AD-A264 298

AFOSR-TR- 93 0289

BAF21
LTR93-003

FINAL TECHNICAL REPORT

4D INTERCONNECT EXPERIMENTAL DEVELOPMENT

29 January 1993

Sponsored by

USAF, Air Force Office of Scientific Research
Bolling AFB, D.C. 20332-6448

Contract No. F49620-91-C-0002

Prepared by
Philip D. Henshaw
Steven A. Lis
Michael G. Cheifetz

"The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Office of Scientific Research."

SPARTA, Inc.
24 Hartwell Avenue
Lexington, MA 02173
Phone: (617) 863-1060
Fax: (617) 861-7934

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	29 January 1993	Final Technical Report 1 Oct 91 - 30 Sep 92

4. TITLE AND SUBTITLE

4-D Interconnect Experimental Development

5. FUNDING NUMBERS

F49620-91-C-002
FQ8671-9001600
3005/AI

6. AUTHOR(S)

Philip D. Henshaw, Steven A. Lis, Michael G. Cheifetz

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

SPARTA, Inc.
24 Hartwell Avenue
Lexington, MA 02173

8. PERFORMING ORGANIZATION REPORT NUMBER

LTR93-003

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

USAF
Air Force Office of Scientific Research NE
Bolling AFB, D.C. 20322-6448

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

3005/A1

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

unlimited

12b. DISTRIBUTION CODE

93-10760

13. ABSTRACT *(Maximum 200 words)*

This report summarizes the work performed during a two-year program aimed at demonstrating the feasibility of constructing a 4-dimensional neural network based on the unique properties of spectral hole buning (SHB) materials. SHB materials were synthesized and excellent quality holograms were recorded and retrieved. Both wavelength and angle multiplexing were demonstrated with no apparent crosstalk. We assembled a demonstration holographic optical neural network and tested it as a bi-directional associative memory system. The results obtained clearly demonstrate the fundamental ability to fully connect two 2D planes of digital information. Expectations are that this architecture can be extended to capacities of 10^12 interconnects or greater.

93 5 18 037

14. SUBJECT TERMS	15. NUMBER OF PAGES
Neural Network, Optical computing, Holographic Storage, Optical Interconnect	41
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	

TABLE OF CONTENTS

LIST OF FIGURES

1 Introduction

This final report provides a description of a two-year program to build and demonstrate the operability of a 4-dimensional neural network computer based upon the special capabilities of a holographically based optical system and Spectral Hole Burning materials (SHB) as the recording media. As described in our Phase I report [1], 4-dimensional capacity is required to fully connect two 2-dimensional planes in a scalable manner. In our architecture, the four dimensions are provided by the three spatial dimensions available using volume holographic recording plus the fourth dimension of laser wavelength. By appropriate system design, one of the input planes can be coded by laser wavelength to make use of this fourth dimension, as shown conceptually in Figure 1.

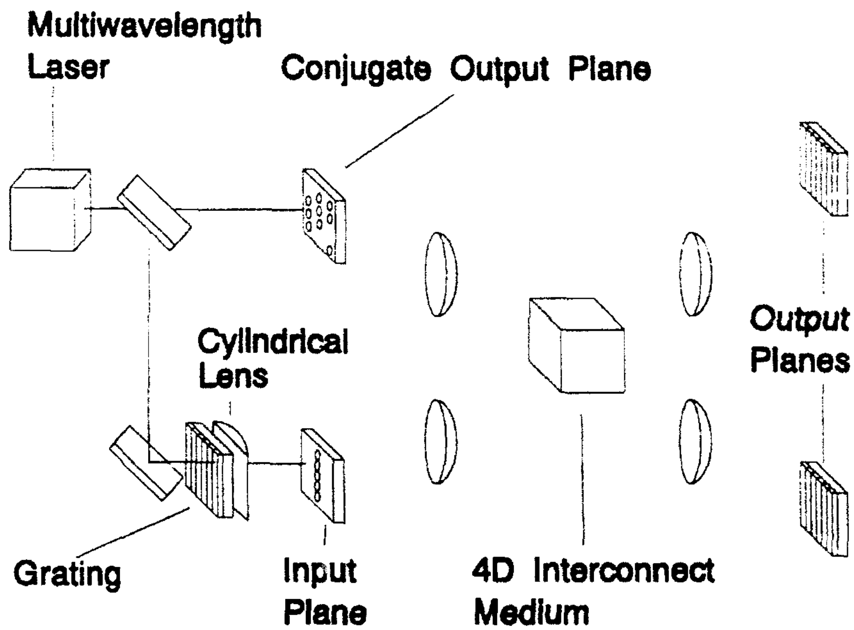


Figure 1. Conceptual diagram of the 4D neural network system based on SHB media.

Our proposed Phase II program was designed to provide experimental confirmation of our Phase I analysis and architecture studies. This experimental program proceeded incrementally from simple experiments to more complex ones – culminating in a complete neural network architecture suitable for demonstration purposes. During the course of our optical neural network demonstration project, we have demonstrated a network with over 1,800,000 interconnections (expandable to over 5,400,000 interconnections). Future versions of this architecture may be able to realize networks as large as 10^{12} interconnections. The successful completion of this program enables us to pursue funding for an upgraded neural network system with over 65,000 neurons, more than 4 billion interconnects, and capable of implementing up to 120 billion interconnects per second.

The Statement of Work for Phase II consisted of three technical tasks plus a management and reporting task. The main body of this report will be organized to discuss each of these items, and the Statement of Work is repeated here for reference.

1. SPARTA will perform a series of experimental measurements to verify the 4D interconnect concept and the analysis performed in Phase I. This experimental program will include the following subtasks:
   a. Spectral hole burning medium selection and preparation
   b. Writing and reading a single interconnect grating
   c. Writing and reading several gratings at different laser frequencies
   d. Writing and reading several gratings at the same laser frequency
2. SPARTA will construct an electronic feedback loop capable of implementing multilayer networks with the capability of learning from training set data. This item will include two subtasks:
   a. Implement a simple electronic feedback system.
   b. Demonstrate that all elements are in-place and working.
3. SPARTA will perform technology transfer planning. This item will consist of three subtasks:
   a. Interview potential government and private end users of neural network technology and determine the applications for which our technology makes sense and why.
   b. Assess the competition from other artificial neural network implementation methods.
   c. Define a possible configuration for a neural network to address one of the potential applications. This configuration should make sense from a performance and cost standpoint.
4. Management and Reporting
   a. Prepare and submit an annual progress report and/or briefings as required.
   b. Prepare and submit a Final Report.

2 Experimental Optical Interconnect Studies

2.1 Preparation of SHB Media

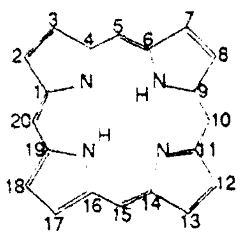
The preparation of SHB media suitable for use as a holographic storage medium has been successful. Polymer matrices based on polystyrene (PS) have been prepared of suitable optical quality and uniformity for our experimental needs. The introduction of the organic photochemical species entailed some subtleties, but has been successful. The information obtained, when coupled with the available literature, provides us with a relatively clear path to the preparation of numerous organic high performance SHB media.

2.1.1 Overview of the Materials Selection Process

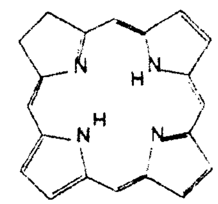
The principal goal of the materials preparation effort associated with this program was to provide a SHB sample of sufficient optical quality and which can be used to record holograms. Since a number of suitable candidates exist which have been previously studied by others, preparation of novel materials was not necessary. Based on arguments which have been stated previously [1] the porphyrins were identified as a class of materials providing suitable performance in terms of homogeneous and inhomogeneous linewidth, absorption cross section and quantum efficiency for conversion. The spectral range at which recording will be performed is determined principally by the selection of the specific porphyrin. Because of the wide range of porphyrins available (see Figure 2) our selection of operating wavelength was made based on optical considerations. For reasons of convenience it was decided to perform the initial studies at 633 nm (the HeNe wavelength). This was decided because of the relative ease of obtaining optical components coated for this wavelength and the convenience of using a simple HeNe laser for the initial experiments. It also turns out that chlorin (H_2Ch) is an SHB material which absorbs strongly in the wavelength range near 633 nm and has very interesting properties for both the near and long term with respect to erasability (which we shall not address here). For this reason, chlorin was selected as the principal material of interest, but we also undertook a small survey of related porphyrins in order to determine the level of effort generally required in the preparation of these materials by different means. To this end, we have also examined phthalocyanine (H_2Pc), porphine (H_2P), chlorin-e$_6$ (H_2Ch-e$_6$), and tetraphenylporphine (TPP).

The host materials which have been previously studied include polystyrene (PS), polymethyl methacrylate (PMMA), polyvinyl alcohol (PVA), polyethylene (PE) and a range of related polymers and copolymers. Most of the previous work by others has focused on the preparation of relatively thin films and often without concern for optical quality because simple absorptive properties were studied. However, high quality holograms have been recorded in thick samples of PS [2] and some significant Japanese work has focused on the used of PMMA and related copolymers prepared from ethylene and methyl methacrylate.[3] Because optical components are often prepared from PS and PMMA (Plexiglass is a common trade name), we chose to center our early investigations on these two host materials.

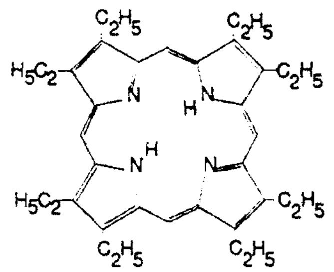
A factor for consideration was the anticipated homogeneous linewidth obtained from the photochemical species in these hosts. Previous workers have shown that these materials have linewidths which are principally determined by the properties of the host polymer.[3,4] The properties of the tunable laser which we used defined the laser linewidth to be roughly 300 MHz and a homogeneous linewidth of similar magnitude is appropriate. Our examination of the



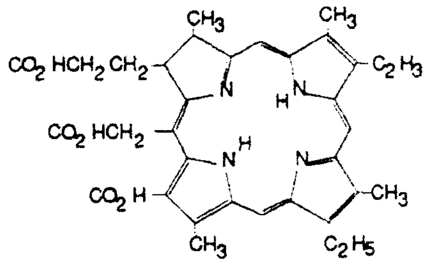
Porphine



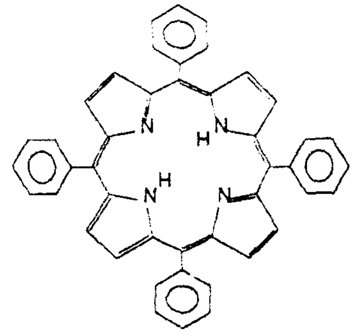
Chlorin I



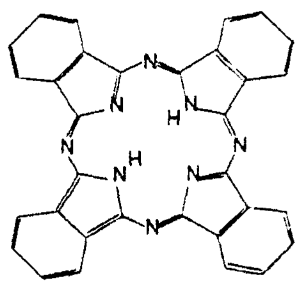
Octaethyl Porphine



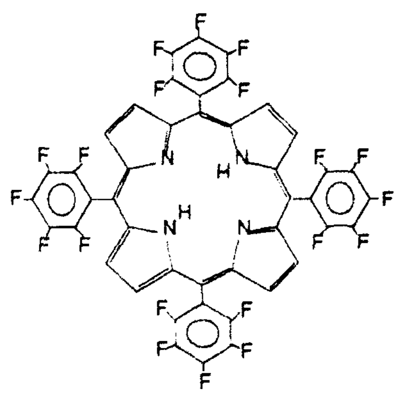
Chlorin-e_6



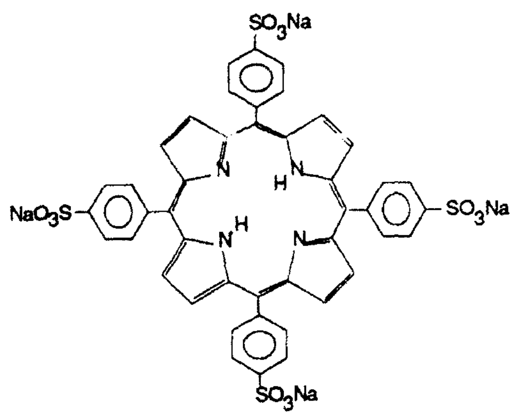
Tetraphenyl Porphine



Phthalocyanine



Tetrakis(pentaflourophenyl) Porphine



Tetrasodium Tetra(sulfonatophenyl) Porphine

Figure 2. Chemical structures of the porphyrin compounds referred to in the literature as candidate SHB materials. Note that in the case of porphine, the numbering system of the carbon atoms has been highlighted. Of these, the first six porphyrins have already been incorporated into suitable polymers by us.

literature indicates that we can expect homogeneous linewidths of the order of 1 to 3 GHz at a temperature of 4 K. for commonly used hosts such as PS and PMMA.

Based on this consideration of properties, we chose H_2Ch in PS or in PMMA for use in our demonstration system.

2.1.2 Properties of the Polymer Hosts

A brief series of simple tests were performed on samples of PS, PMMA, and PVA (already polymerized) purchased from chemical suppliers.

1. It was quickly determined that PS will melt easily at 150 C and can be readily handled and cast into a number of shapes. Heating in air produced a modest odor and also resulted in a faint yellowing of the polymer (this was indicated to be likely from comments found in the literature [5]). We found that melting the polymer under a nitrogen atmosphere provided excellent results.
2. PMMA would not melt, but would tend to decompose and rapidly vaporize (rather cleanly but with an unpleasant odor). It was clear that the simple melting procedure used for PS could not be used here, but that direct polymerization of methyl methacrylate (MMA) would be required.
3. PVA would not melt, and on decomposition would form a tan to black residue which was rather unattractive.

All the above polymers would readily dissolve in the right solvents, but it generally required a fair amount of solvent to produce a solution of modest viscosity. The eventual goal was to produce a relatively thick sample, free of solvent, so removal of the solvent was quickly seen as a problem. Attempts to drive off the solvent from samples a few mm thick lead to severe bubbling which results in samples having unacceptable optical quality. Based upon our need for relatively thick samples, it was quickly determined that the samples must be either melted and cast (as in the case of PS), or directly polymerized into the shapes desired. Optical polishing of plastic surfaces proved to be a significant problem and we found that simple casting of the PS samples provided excellent quality samples with minimal effort.

Based on the above considerations, it was concluded that attention should be directed to the melting and casting of PS for use as a host material.

2.1.2.1 Free Radical Polymerization

The polymerization of styrene and MMA is quite simple with the only negative aspect being the odors generated. These materials are of modest volatility and should only be handled in a ventilated room. The principal means for polymerization of these two materials is via free radical mechanisms. The free radicals can be readily provided by chemical additives called "initiators". The free radicals are generally provided by the breakdown of the initiator molecules.[6] This breakdown can be driven thermally (as in the case of benzoyl peroxide) or via UV illumination (as in the case of benzoin). Once the free radicals are generated, they become the "seeds" for creating long polymer chains. When the free radical adds to the monomer molecule, the new molecule retains the reactive capacity to add to more monomers (see Figure 3). This process continues until the radical couples to another radical, ending the chain. Via this chain growth mechanism, a rather small amount of initiator (0.5 percent or less) is all that is required to provide full polymerization.

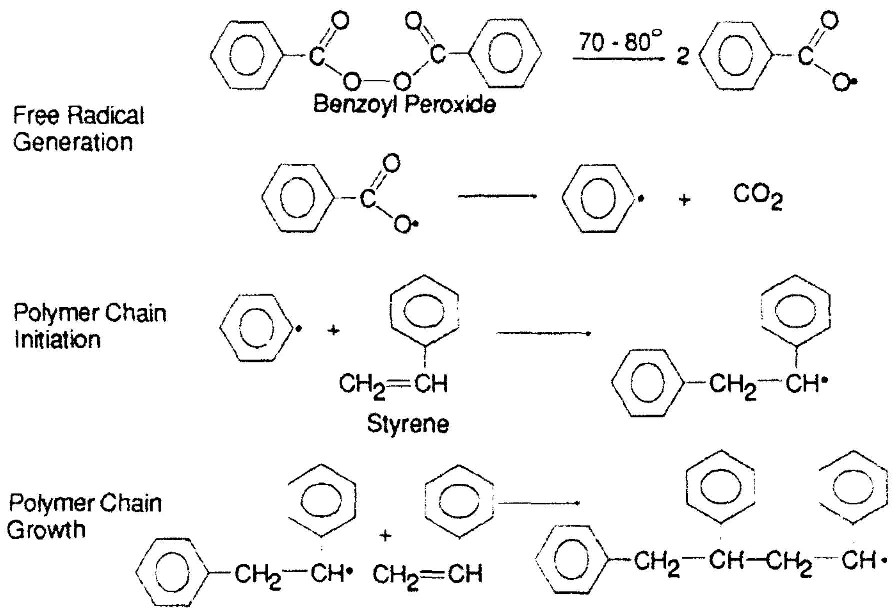


Figure 3. This is an example of the free radical polymerization of styrene which is initiated by the thermal decomposition of benzoyl peroxide. Once the polymer chain growth has been initiated, it can continue until one of several chain termination events occurs.

2.1.2.2 Polymerization of MMA Using Benzoyl Peroxide

Benzoyl peroxide readily decomposes at 70 to 80 C to form the necessary free radicals. It is readily soluble in MMA. Our procedure for preparation generally involved placing the MMA/benzoyl peroxide solution in an aluminum weighing dish, which is then placed in a glass petri dish (with cover). Upon heating to 80 C in an oven, polymerization began in about 1 hour, and was fairly complete in 2 hours. Care needs to be exercised in not trying to polymerize a sample which is too thick because thick samples can bubble severely. Bubbling can be a problem during the polymerization process because the polymerization reaction is quite exothermic. In principle this could be handled by attempting to control the temperature. however, this is quite difficult. The cause has been suggested in the literature [7] to be related to the steric hindrance provided by increased viscosity. It is believed that when the MMA is approximately 40 percent polymerized, molecular motions become adequately inhibited so that ending of the chain reaction via the combination of two radical chains becomes unlikely. The chains then continue to grow rapidly, and the exothermic reaction has been reported to be explosive in certain cases.

We have found that this problem can be adequately avoided by polymerizing samples which are only 2 to 3 mm thick at a time. After polymerization (with some loss of MMA due to vaporization) one can obtain samples 1 to 2 mm thick. Successive layers of PMMA can be easily built up by adding more solution, and a sample of 1 cm thickness can be obtained with a few hours more work. The general optical quality is good, although the total absence of bubbles and uniformity of thickness is achieved when the amount of care taken is increased.

2.1.2.3 Polymerization Using Benzoin

A commonly used alternative to the initiation being driven thermally is to use a photoinitiation process. This has the advantage of permitting one to carry out the polymerization at a lower temperature. The lower temperature is an advantage because of the reduced odor problem, a reduced MMA loss due to vaporization, and a reduced problem associated with the bubbling caused by the runaway exothermic reaction.

A typical procedure involves preparing a solution of MMA and benzoin (0.5 percent). The solution is again put in an aluminum weighing dish and inside a glass petri dish (with cover). A low power UV lamp is then placed over the sample. The initial increase in viscosity takes 3 to 6 hours and complete polymerization takes 6 to 18 hours. While being a slower process it also seems to be a much lower temperature process. The heating by the lamp does not raise the sample temperature above 40 C. Samples 5 to 6 mm thick can be obtained in a single step which are of high optical quality and completely free of bubbles. (Some bubbling was noted for a sample 1 cm thick.) With the modest number of samples prepared so far, it can be surmised that the sample quality provided by the UV photoinitiation procedure is higher and less work is involved.

2.1.3 Adding the SHB Species

Several considerations are of importance in introducing the SHB molecules. First is the relatively limited solubility of these materials. This has proven to not be a problem, since the optical density obtained is adequately high due to the high molecular absorption cross section. Another factor for consideration is that the best results in terms of the performance of the SHB materials can be expected if the use of solvents is either completely avoided or at least minimized.[4] Therefore, direct dissolution in the polymerized material or monomer is desired.

2.1.3.1 Polystyrene Samples

In the case of the materials selected for examination, all, except for H_2Ch-e_6, were directly soluble in polystyrene. This approach to preparation was tedious, but did not involve an intervening solvent. In each case, the solubility limit seemed to coincide with the absorption reaching an optical density of about 1 in a 1 cm thickness. Dissolution was carried out by simply weighing out a few milligrams of the SHB species and stirring it into a melted sample of polystyrene. Repeated patient heating and stirring was required to achieve complete dissolution, and dilution through the addition of more polystyrene was often required to finally eliminate all the visible solid porphyrin material. It became clear that diffusion of the SHB molecules through the PS was an important process to enhance the dissolution, and it appeared that the H_2P had the most rapid dissolution, while the H_2Ch-e_6 appeared to be soluble, but was not able to diffuse rapidly. All the others appeared to be amenable to this rather direct method for introduction into the PS.

The problem with H_2Ch-e_6 was overcome by initial dissolution of the material in 1 ml of dimethylformamide. This solution could then be added directly to a large amount of polystyrene to provide appropriate dilution. This procedure provides samples with good clarity and uniformity and with an absence of particulates.

Realizing that prior dissolution in a small amount of solvent could eliminate the particulate issue, we found that H_2Ch-I was readily soluble in a few drops of dichloromethane which could then be added to the polystyrene. Upon melting of the polystyrene, most of the solvent vaporized and the H_2Ch-I readily dissolved in the PS providing samples with excellent optical clarity and quite high optical densities if desired. We therefore concluded that the use of a small amount of solvent was required.

2.1.3.2 Polymethyl Methacrylate Samples

As was stated above, the preparation of the PMMA samples was not possible by simple melting of the polymer, and the preparation of thick samples, free of solvent, from solution could be tedious. The principal path of interest then lies in the direct polymerization of the MMA solution which contains the SHB material. It was found that this approach does not permit all porphyrin compounds to be used in such a procedure.

The limitation was found to be caused by the attack of the free radicals on some of the porphyrin compounds. If the porphyrin ring is broken by free radical addition, the dye becomes effectively bleached. It was found that this process could be avoided by proper selection of the chemical groups added to the central porphyrin ring and their exact locations. The reasoning behind why this happens is presented in the discussion which follows.

MMA undergoes free radical addition across the double bond between two carbon atoms. Such an addition can occur across any unsaturated bond (in principle), however, unsaturated bonds which are highly stabilized due to resonance effects (as in benzene) may not be appreciably attacked. A simple example of this is styrene which undergoes free radical addition across the lone C-C double bond (to form polystyrene), but attack does not occur on the accompanying phenyl group (see Figure 3).

In the case of the porphyrin ring compounds one may expect that free radical attack may be more probable at certain locations on the exterior carbon atoms of the central ring. If additional chemical groups can be added to the carbon atoms most available for attack, we find that bleaching of the SHB species is now inhibited. A strong suggestion that this was possible was made by several Japanese references.[8,9,10] In particular these references indicate that direct polymerization of MMA and related monomers is possible without bleaching of the porphyrin compound if the porphyrin ring is surrounded by phenyl groups attached at locations 5, 10, 15, and 20 on the porphyrin ring (see Figure 2) as in the case of tetraphenyl porphine (TPP). It is believed that these phenyl groups provide hindrance to the attachment of free radicals simply due to their physical size. There is also an indication from the literature that the added phenyl groups are not restrained in their orientations relative to the central porphyrin ring, rather, they are able to rotate freely in solution.

The exact location of the additional groups at locations 5, 10, 15, and 20 also seems to be important. By comparison, we determined experimentally that ethyl groups added to locations 2, 3, 7, 8, 13, 14, 18, and 19 (as in the case of octaethyl porphine) did not provide any significant protection against free radical attack because this compound was found to bleach readily under polymerizing conditions. The problem of bleaching was also shown to not be simply due to the presence of the initiator, since a solution of porphine in MMA (without any initiator added) bleached in a few hours without any noticeable polymerization taking place.

It would appear that unless the porphyrin ring is stabilized by the protection of locations 5, 10, 15, and 20, the introduction of these materials into polymers cannot be carried out under conditions where free radical polymerization is possible. This turns out, however, to not be a severe restriction because of the large number of related compounds which are still available for use. It should also be noted that the key aspect which distinguishes the chlorin derivatives from the porphine derivatives lies in the saturation of the bond between locations 2 and 3. Since these locations do not seem to be key to preventing free radical attack, we can expect that the chlorin derivatives of the tetraphenyl porphines are also stable against bleaching during polymerization.

The conditions under which PMMA could be directly polymerized and not cause bleaching of the SHB material were not studied extensively because it was not warranted under this program. The results, however, were deemed to be very encouraging in that all samples necessary for carrying out this program were shown to be easily prepared; a clear path exists for future preparations of materials of higher performance (homogeneous linewidths smaller than 1 GHz).

It was shown that TPP doped samples of PMMA could be readily prepared as thick slabs (3 to 10 mm thick) via direct polymerization. It was demonstrated that no noticeable bleaching occurred whether the initiation process was driven thermally using benzoyl peroxide or via UV light using benzoin. Both procedures worked well, with the benzoin providing samples of somewhat better optical quality. The success of the benzoin was somewhat surprising because of the strongly absorptive properties of the TPP itself in the UV. We were concerned that the TPP might prevent sufficient light from reaching the benzoin initiator molecules and prevent efficient polymerization. It would appear that this is not a problem, and that the TPP is very stable under these long term UV exposure conditions. (The success of photoinitiated polymerization also turns out to be strong evidence supporting the contention that the SHB medium will **not** be sensitive to **irreversible** photobleaching. One can therefore expect that these materials will permit a long product life.)

A simple test was performed using crossed polarizers which indicated that the PMMA samples prepared were relatively stress free. However, a problem arose concerning the ability to provide samples with optically flat surfaces coupled and which were stress free. We previously noted that polishing of plastic surfaces was difficult. We could obtain flat surfaces through casting, but the polymerization process was difficult to control using our methods. The net result was that the polymerization process could result in thermal stresses which would result in stress birefringence in the optical samples. While casting procedures have been developed by the chemical industry for providing PMMA of suitable optical quality (both surface quality and uniformity) our limited facilities do not permit us to duplicate their methods. The most common approach used by the plastics industry for PMMA processing is high pressure injection molding. While this process requires special facilities, if is widely known to provide plastic components of suitable optical quality for demanding applications such as lenses for optical disk readers.

Similar tests of the PS samples indicated that stress free material could also be prepared if care was taken to cool the samples slowly. Since essentially stress free samples could be easily prepared, from PS which had excellent optical surfaces we chose to single out PS as the optimal host medium for this program.

2.1.4 Indications of Future Materials Development Paths

For the preparation of uniform thick samples of doped polymers having arbitrary composition, direct polymerization would appear to represent the most clear path to obtaining high quality samples. This approach is favored in the long term because while PS can be easily melted and mixed with many SHB materials, the smallest homogeneous linewidths have been reported for polyethylene (PE) and copolymers of ethylene and MMA.[4,3] Also, while PMMA is optically very clear, PE is generally translucent. This translucence is the result of the tendency of PE to form polymer crystals upon cooling which scatter light. Very rapid quenching of melted samples can inhibit this process somewhat, but the long term stability of such a sample may be in doubt.

The interest in the performance of PE is caused by the observation that when used as a host material for the SHB species it provides the most narrow homogeneous linewidth observed among the organic SHB media. In general the homogeneous linewidths of porphyrin dopants in PMMA or PS are roughly 20 times larger than that observed for the same materials in PE. The solution to this divergence between optical clarity and homogeneous linewidth is provided by Japanese workers [3] who have shown that the observed homogeneous linewidth is linearly related to the ratio of components used in preparing copolymers of ethylene and MMA. Their experimental results are plotted as a function of number of carbon atoms in the alkyl chains in Figure 4 and indicates the ability to prepare copolymer samples of virtually any copolymer ratio in an attempt to obtain the homogeneous linewidth desired. One would also expect that at some well defined copolymer ratio the samples will be of adequate optical clarity for our purposes. This expectation is based on the realization that one simply needs to prepare copolymers where the local structure is adequately irregular to prevent crystallization. This can be provided by the preparation of ethylene/MMA copolymers or by the polymerization of alkyl methacrylate copolymers where the alkyl groups are rather long (see Figure 5). In both cases one find that the homogeneous linewidth approaches that found in PE.

The selection of the porphyrin compound determines the wavelength range over which hologram recording can be accomplished. H_2P (the simplest of the porphine derivatives) absorbs with its 0-0 vibrational band at 611 nm. H_2Ch (the chlorin analog) has a very well-defined strong absorption at 635 nm. In general, for every porphine derivative, there is a matching chlorin derivative. The absorption wavelength of the chlorin derivative is generally located 25 nm longer in wavelength. The addition of chemical groups at any location in the porphyrin ring system generally increases the absorption wavelength. For example, TPP has an absorption peak at 647 nm, while octaethyl porphine absorbs at 618 nm.

The addition of chemical groups at locations 5, 10, 15, and 20 on the porphyrin ring does not seem to be a factor which significantly restricts the material selection. In fact, the process for preparing these porphine derivatives is easier than that for preparing the basic porphine material itself. This relative ease in preparation is reflected in the significantly lower prices on the porphyrin derivatives ($10/gram) having several additional chemical groups compared with the simple H_2P and H_2Ch compounds which are far more expensive ($100/milligram). Introducing modifications to the phenyl side groups alters the absorption peaks further. For example, tetrakis(pentafluorophenyl) porphine absorbs at 659 nm, while tetrasodium tetra(sulfonatophenyl) porphine absorbs at 640 nm. As we can see, the modest requirement that phenyl groups be attached to the porphyrin ring at specific locations is not a serious limit, since a range of SHB species are available and all those mentioned above have been reported in the literature as SHB recording media.

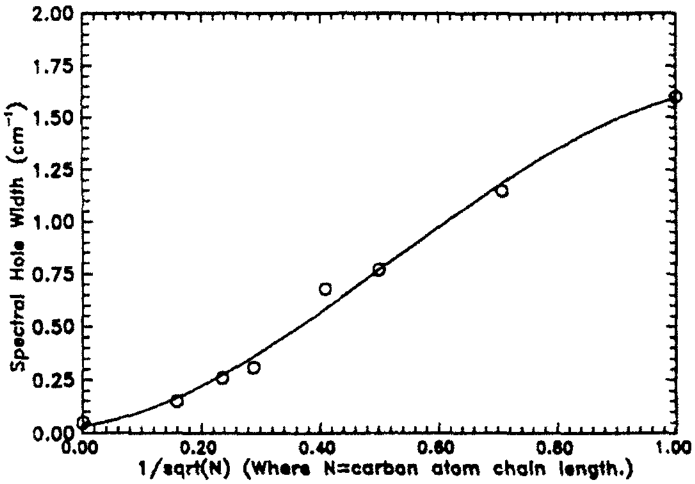


Figure 4. Spectral hole width for 1,4-dihydroxy-9,10-anthraquinone in various polymer host matrices. (This material was selected as representative of SHB media in general.) Note the simple correlation between hole wiath and polymer composition.

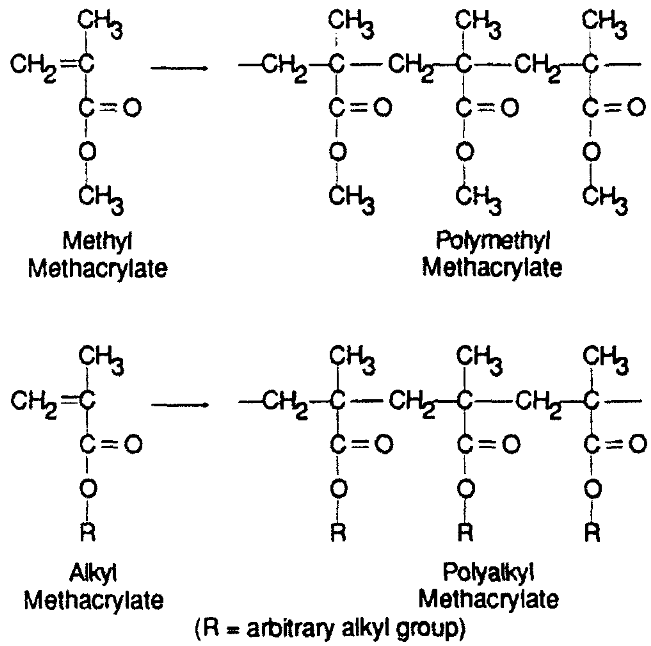


Figure 5. This diagram shows how free radical addition across the C-C double bond in MMA permits polymer chain growth. By simple substitution of any number of selected alkyl groups for the methyl groups present one can alter the alkyl content of the polymer and thereby control the homogeneous linewidth.

The same variety of compounds exists for the analogous chlorin derivatives with the same trend in peak wavelengths. An advantage to the chlorin analogs is that the spectra are generally simpler in structure, with each compound having only a single strong absorption peak in the red or near IR corresponding to the 0-0 vibrational band. This presents an advantage when considering the potential use of a number of photochemical species in the same polymer, for the spectra will not strongly overlap and interfere with each other. By contrast, the spectral structure of the porphines can be far more irregular, with several shorter wavelength peaks having strong absorptions which are not well spaced from the principal long wavelength absorption.[11] All things considered, the chlorin derivatives, which also contain the chemical structure inhibiting free radical attack, should be rather easy to obtain and simple to introduce into polymers, providing the desired spectral properties.

In the long term, an extra concern for data longevity must be added to the SHB material requirements. This has been identified as having a straightforward solution through the use of deuterated host materials.[12] By replacing the hydrogen with deuterium in the host medium, the data life can be extended from weeks to thousands of years. Given the ready availability of deuterated organic chemicals, and the relatively small quantities of polymer host material actually required for each computer system, this should not be a significant technical or cost factor for the future.

2.1.5 Summary of Media Preparation

The materials preparation carried out so far has been very successful. Simple procedures have been worked out which provide polystyrene samples of high optical quality. It has shown that all of the porphyrin compounds examined are readily compatible with our polystyrene preparation method which is a simple melting and casting process.

When the preparation of PMMA or other copolymer mixtures is desired, direct polymerization via free radical initiation is the preferred method of preparation. UV photoinitiated polymerization seems to be the superior method using simple procedures. The polymerization process indicates that a specific class of porphyrin derivatives should be used in order to minimize bleaching through free radical attack.

It must also be remembered that the free radical initiation mechanism is only one of several major approaches available for the preparation of polymers. The specifics for the free radical polymerization of MMA and styrene have been presented and used in this program only as examples. While the results obtained are excellent, the capability to provide SHB material via numerous alternative approaches also exists.

Based on the wide range of porphyrin compounds available, the range of literature information available, and the general flexibility inherent in the ability and ease with which polymers can be prepared, a clear path can be seen to the preparation of SHB materials for both modest and high performance requirements. Future materials preparation efforts will need to rely more heavily on the vast experience of the plastics industry for the preparation of optical samples.

2.2 Optical System Assembly

The optical system design followed very closely the design initially outlined in our Phase I report and proposal. The details provided in those documents will not be repeated here. The general optical layout is shown in Figure 6 and should be compared with the schematic diagram of the system seen in the Introduction. The principal features of the system which have been refined are;

1. the tunable laser system,
2. the camera selection, and
3. the LCD selection.

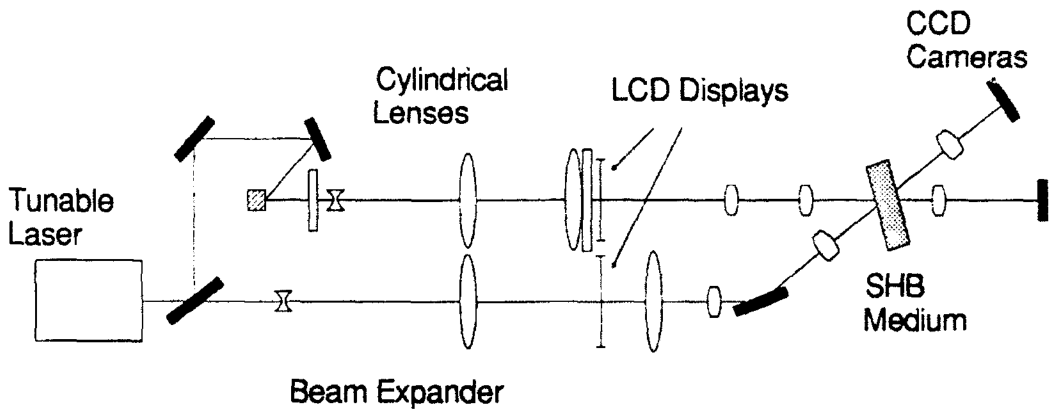


Figure 6. Detailed layout of the optical system which has been assembled. With the exception of the LCDs the key components have not been significantly altered from the originally proposed design.

2.2.1 The Tunable Laser System

The requirements for the tunable laser are that it provide a set of equally spaced wavelengths which can be accessed via computer control. It would be desirable to have at least several mW of power and the laser linewidth should be comparable to the homogeneous linewidth of the SHB medium selected. The tuning rate need only permit scans of the laser wavelengths in seconds as opposed to milliseconds. Finally, the tuning must be repeatable over the duration of our experiments (hours).

It had originally been suggested that this program would utilize a Frequency Agile Laser (FAL) which was electro-optically tunable and readily available at SPARTA as our laser source for the Neural Network Optical System. However, our analysis of our performance needs indicated that a more correct choice would be a tunable dye laser system which is tunable through the use of a mechanically adjustable birefringent filter (BRF). While the dye laser based FAL system could provide excellent speed in terms of wavelength random access, we found that the laser linewidth was in the many gigahertz range.

The advantage of the BRF tuner was that it provided a fundamentally narrower laser line as well as having fewer losses. The lower loss level provided us the ability to introduce etalons as additional tuning and line-narrowing elements. Our composite dye laser design made use of a BRF tuner, a 0.5 mm thick etalon and a 10 mm thick etalon. The composite performance resulted in a laser linewidth of roughly 0.3 GHz. Moreover, by simple rotation of the BRF tuner one could change wavelength by approximately 300 GHz increments. In our tests of hologram recording (to be discussed later in this section) we have found that chlorin-doped polystyrene provides a range of operation of approximately 10 nm. Over this range we could achieve a suitably constant laser line width at a set of wavelengths by simultaneously rotating the BRF filter and tilting the etalon. The rotation and tilt values to access any laser line used for hologram recording were stored in a lookup table which was used for computer control of the linewidth.

The tuning of the birefringent filter is provided by an Oriel "Encoder Mike," a motor driven micrometer system which is computer controllable through a standard RS-232 port from a PC. This micrometer provides adequate precision for reliable "hopping" between laser lines in a repeatable manner. A modest amount of care must be exercised for positioning of the BRF tuner to avoid laser emission at two laser lines simultaneously, however, this is easily accomplished through positioning precision on the level of one percent.

2.2.2 Camera Selection

Early in the program, it was found that a CCD camera manufactured by Electrim Inc., was of a format, size and performance quite ideal for our applications. This thermo-electrically cooled camera has a modest pixel format (192 by 165) which matched well our choice for the input spatial light modulators. The camera image is automatically digitized to 8 bits and read directly into PC memory. This conveniently places the camera image directly at the disposal of any high level programs for further analysis of the image and extraction of digital data from the holographic images. The software to control the cameras is available as C linkable routines which can be readily incorporated in the overall system software. Furthermore, multiple cameras can be operated from the same PC and are completely software selectable and controllable in all the critical control attributes.

2.2.3 LCD selection

The selection of the Liquid Crystal Display (LCD) was a key item in terms of data format, image quality and convenient match with the cameras. Our experience in a previous holographic memory program has shown that commercially available LCDs are quite compatible with the requirements of our demonstration system, however the performance was only marginal when considered in terms of image contrast. The LCD which had been used in the previous program was not an active matrix display, therefore it suffered from two problems associated with contrast. First, the contrast was limited to roughly 20:1, which we found to cause a noticeable loss in image quality. We found that the active matrix arrays, which boast a contrast of 100:1, enhance image quality and reduce noise in the resultant recorded digital data. Second, the array we had previously selected demonstrated a "shadowing" problem which adversely affected image resolution. This shadowing problem was apparent to the naked eye when a single pixel was turned "ON". The adjacent pixel, which was addressed by the LCD sequentially, became partially turned on, resulting in a shadowing effect. The net result was to reduce the effective resolution of the LCD in one dimension by a factor of two. We found that an active matrix LCD provided better performance on these two points because of the greater independence of the transistorized nature of each pixel.

With the requirement added that the LCD be an active matrix one, the composite requirements of the LCD can be summarized as:

1. active matrix (for greater contrast),
2. black and white with gray scale (a color LCD pixel is usually composed of 3 adjacent LCD pixels with integrally incorporated color filters which would cause significant light loss in our application),
3. readily addressable through a commonly available PC video interface (although even composite video could be adequate for our needs),
4. LCD must be able to be illuminated transmissively, and
5. the pixel dimensions (aspect ratio) should permit a good match with the cameras selected.

These requirements were well met by the performance of a liquid crystal projector system from Sharp, the XG-1500U. This system utilizes three separate LCD screens, the black and white images from each being coupled to produce the full color images. (Each LCD is transmissively illuminated with chromatically filtered light, and then the three primary color images are recombined as a single optical image.) The system provides the computer interface required, and is a high contrast active matrix system. An extra benefit is that a single projector system provides three separate LCD screens of excellent quality, a factor which actually makes the system quite a cost-effective choice since two LCDs were required for the program.

2.3 Holographic Test Results

During the optical system testing as well as the material preparation we performed a number of holographic tests to determine performance with respect to the optical recording parameters. Since the material properties are generally known through the available literature, our results could be readily compared to verify whether our optical and cryogenic systems as well as the material were performing as anticipated.

2.3.1 SHB Medium Cooling

The holographic tests performed verified that the spectral hole burning medium selected (Chlorin-I in polystyrene) was performing as anticipated. It was generally observed that the preparation of the material required some care on a few important items in order to provide the best possible image quality. We identified three key items:

1. care must be exercised to prepare the material under circumstances which are free of any particulate matter including dust and insoluble residues,
2. the host polymer must be cast into the proper shape required for the cryogenic system, however, this must be accomplished in a manner which induces no residual stresses or non-uniformities in the material, and

3. because the host polymers (polystyrene in particular) are stress birefringent, the samples must be held in the cryostat in a manner which applies minimal force to the sample, however good thermal contact to the sample must be maintained to keep the material at cryogenic temperatures.

The first two items above are simply a matter of adopting suitable procedures for sample preparation the details of which are not particularly interesting so they will not be discussed in detail here. The third item (stress free thermal contact in the cryostat) has provided us with some cause for concern, but an excellent solution was adopted which has positive short-term and long-term ramifications.

The cryostat which we had purchased for this project was a compact "Supertran" cryostat manufactured by Janis Research. This cryostat was selected because of its low cost as well as the fact that the basic structure of the "cold finger" resembled that of the commercially available refrigerator systems. Inherent in the design is the concept that the sample of interest is cooled conductively by mechanically making good thermal contact to the sample. We found that this approach caused serious optical distortions of the recorded images and our measurements suggested that the sample temperature was not at the low temperature of the sample holder.

A common alternative approach to sample cooling at these temperatures (particularly for samples which conduct heat poorly) is to surround them in a helium atmosphere, which is highly thermally conductive. This approach generally uses the same helium supply to provide total cooling as well as thermal contact. Unfortunately, we saw this approach as a long-term limitation when considering the use of a closed cycle refrigeration system. Our solution to this issue was to fabricate a sample holder which incorporated a static helium atmosphere surrounding the sample while the primary cooling is provided via mechanical contact to the cold finger assembly (see Figure 7). This sample holder provides excellent thermal contact to the polystyrene sample. Optical transmission is permitted via the sapphire windows, and the flexible vacuum seals are of indium metal. Thermal conduction leakage to the sample holder via the helium gas tubing is minimized by making the tubing of stainless steel. While this design permitted us to avoid the obvious extra expense of purchasing a new cryostat, it also verified a sample holder design which can easily be engineered to be compatible with a closed cycle cryostat system.

2.3.2 Standard Holographic Tests

Fundamental to the operation of the neural network system is the ability to record and read out high resolution holograms from both input planes. Furthermore, we are required to demonstrate that both angle multiplexing and wavelength multiplexing are readily possible utilizing our system design. We successfully confirmed these capabilities during some of our early tests. These tests were carried out by placing a chrome mask of the Air Force resolution chart in the 2D image plane (see Figures 8 and 9). To create a suitable reference beam, a slit was inserted in the chromatically steered leg of the system which, because the illumination is brought to focus as a single line, causes the illumination to be reduced to a single point. The light from that single diffraction-limited point then uniformly illuminates the SHB medium for proper interference with image containing beam. Lenses and cameras are placed in the beams after the light traverses the SHB medium, and the images and holograms can be viewed directly via the CCD cameras.

Typical results are shown in the figures which follow. These CCD camera images are comparisons of the input images and the holographically retrieved images as viewed by each of

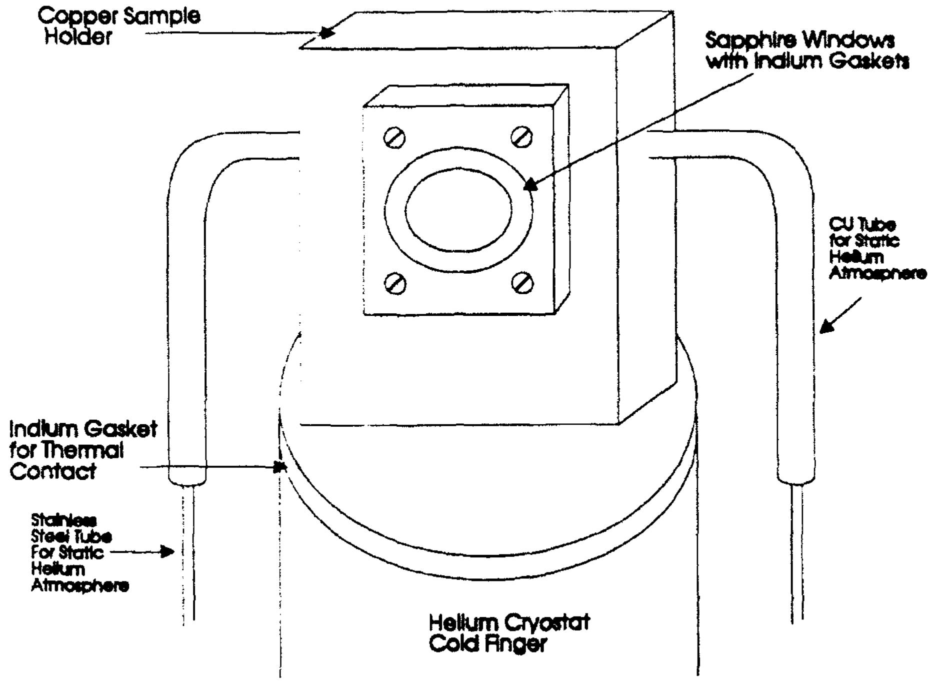


Figure 7. Diagram of the sample holder design which permits a static helium atmosphere for stress-free thermal contact to the SHB medium. The cooling is provided conductively via mechanical mounting of the sample holder to the cold finger of the cryostat.

the CCD cameras. Holograms from both legs can be viewed from a single recorded pattern. As can be seen from the images viewed below, the input and output resolution is nearly identical. The approximate hologram efficiency observed was roughly 10^{-3} and the exposures required to create these holograms closely matched our estimates based on literature reports of material properties. Some background scatter is seen in the holograms which we attribute principally to residual particulate matter in the polystyrene sample and accumulated dust on some of the optical components. We have shown that a modest increase in the care of sample preparation and a cleaning of the optical elements provides even clearer holographic images.

Simple tests were also performed which showed that both angle multiplexing and wavelength multiplexing are readily possible for the recording of multiple holograms. This capability was clearly demonstrated by angle multiplexing three holograms and wavelength multiplexing 10 holograms simultaneously with no apparent crosstalk. The separation in angle and wavelength matched our theoretical and design requirements very precisely. This cursory test was not at all limited by system performance but only by experimental convenience.

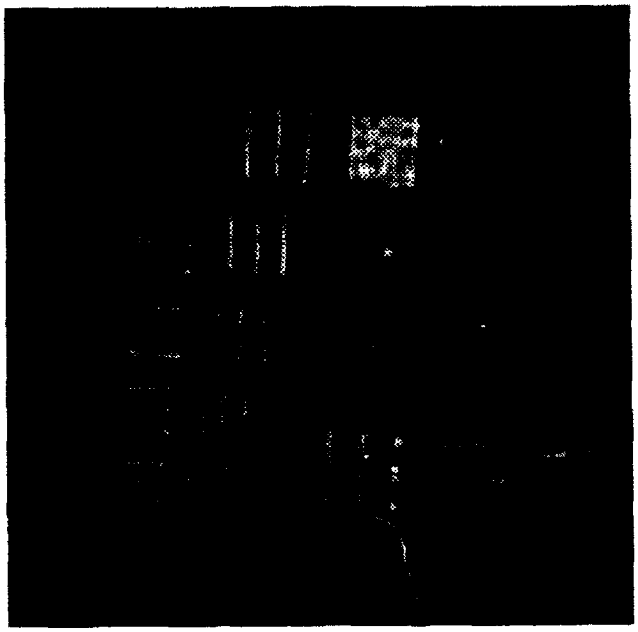


Figure 8. Input image of the Air Force resolution target on a chrome mask used to test the resolution of the optical system and the CCD cameras.

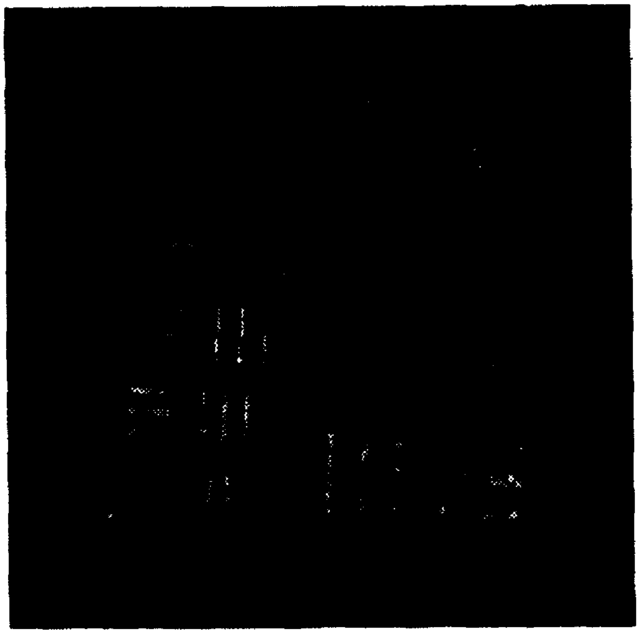


Figure 9. Recorded hologram of the Air Force resolution target. Note that the resolution of the image is principally limited by the CCD camera and the minor imaging imperfections can be traced to imperfections (dirt) on the optical components.

3 Demonstration – Integration of Optical and Electronic Subsystems

3.1 Overview of the Demonstration

In this section, we will describe the integration of the optical interconnect hardware described in the previous section with a PC-based system for input, laser control, optical detection, and non-linear pixel by pixel processing to implement complete neural network operation. In order to demonstrate that this system showed full neural network functionality, we operated it as a bidirectional associative memory (BAM). We were able to associate two images, "SPARTA" and "AFOSR", showing that one image could be recovered given the other, and we were able to recover the complete "SPARTA" pattern given only "PART" of SPARTA.

In the process of developing the software and testing the network operation, we encountered problems which were solved by two techniques: slope coding and phase coding. These two techniques have important implications for memory as well as neural network systems; for this reason we describe them in some detail here.

It must be noted that the two SLM planes are actually architecturally symmetric. From the users point of view, the data encoded in each may be presented in identical manners and without concern for fractal spacing. Each plane can be fully populated. Learning and calculating can be carried out either via illumination at one laser wavelength at a time and sequentially stepping through all the wavelengths corresponding to each of the rows of the input plane or via illumination by all of the wavelengths simultaneously.

The utilization of two cameras as well as two SLMs permits the system to operate as a completely general neural network capable of carrying out the functions of learning, interconnection, feedback and backpropagation with complete symmetry.

3.2 Experimental Implementation

A schematic diagram of the optical system we have constructed is shown in Figure 10. The entire system is composed of commercially available hardware components which have been described in the previous section. The entire system (laser, shutters, LCDs and cameras) is under direct control of a single PC compatible computer with appropriate interface boards. All system control software is written in C or was available as commercially available software modules.

The SHB medium used is chlorin-doped polystyrene which has an absorption peak at 635 nm. The thickness of the medium was selected to provide efficient Bragg angle selectivity so that independent LCD pixels could be accessed and recorded holographically. The results presented here utilized a spectral range of approximately 7 nm of the 10 nm wide inhomogeneous absorption band.

For efficient use of the CCD cameras, a direct correspondence must be made between the input SLM pixels and the output CCD camera pixels so as to permit rapid and efficient usage of the cameras and software. This was accomplished by careful adjustment of the system magnification add the addition of small cylindrical lens elements which permitted modest changes to the aspect ratio to be made for optimal match.

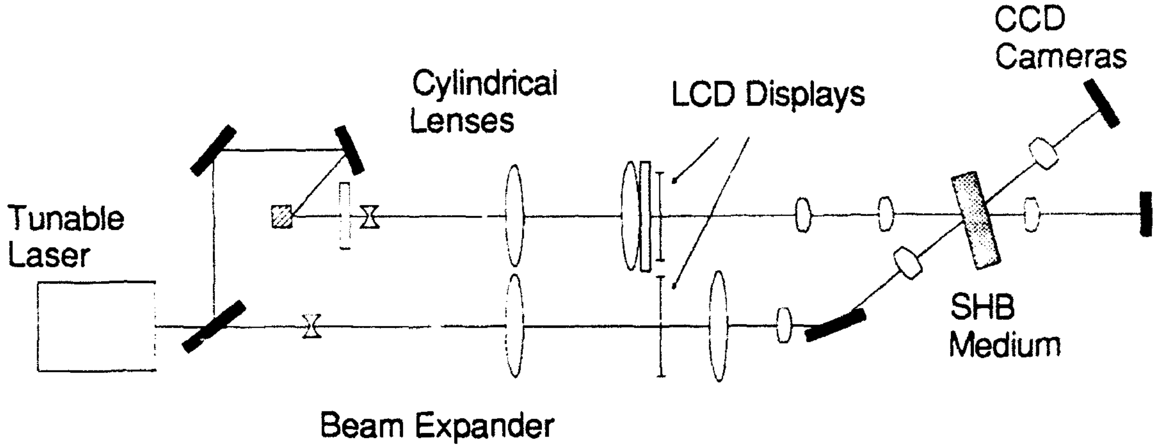


Figure 10. Schematic of the 4D interconnect experimental system. Polarizer components, shutters, and the cryostat are not shown.

3.2.1 Slope coding

Error correcting modulation, or slope coding as we call it, has been discussed by Howe. *et al.*[13] Slope coding reduces errors by using signal patterns with characteristics that are well-suited to the recording medium and is often used in high performance magnetic and optical disk media. In this approach, a single bit is recorded as a local modulation in the recording medium characteristics so that an increase in signal may be interpreted as a 0 and a decreasing signal might be a 1. This makes the system only sensitive to local modulation of material properties and less to large scale variations.

Variable signal strength is a recognized problem for holographic systems, in part because of erasure during subsequent recording and readout operations and also because of illumination nonuniformities. The effects of variable signal strength can be eliminated to a large degree using slope detection. Using a CCD detector array, each data value requires two pixels, resulting in 100% overhead, however it markedly reduces the raw bit error rate of the system.

For the neural network experiments reported here we have chosen to adopt a trinary system of data coding as shown in Figure 11. In this scheme our experimental images are presented as 1's on a background of 0's. Negative pixel values were not input but were used as potential output values in this experiment. This has the advantage of enhancing image identification of the neural network.[14] A simple threshold easily defines the cutoff between interconnect values.

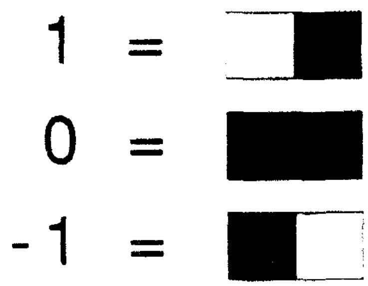


Figure 11. Examples showing how the trinary information is encoded.

3.2.2 Phase coding

For proper operation of the neural network, the SHB medium should be located at or near the Fourier transform plane of both optical legs of the system. This permits proper connection between every pixel in one plane and every pixel in the second plane. To avoid hot spots in the recording medium which could lead to nonlinear recording properties, we have implemented a random phase coding procedure which presets the polarization of each pixel in each 2D input plane. The pattern of UP and DOWN polarizations is randomly determined but identical for all patterns presented to the system. This scheme has two principal beneficial effects. (1) The light intensity at the SHB medium is well distributed (without significant light loss) independent of the regularity of the input pattern. (2) The phase code provides a mask of coding which permits one to discriminate between inputs which have been shifted vertically thereby re-enforcing the unique identity of each pixel.

The combination of slope coding and phase coding requires simultaneous amplitude and phase modulation. This dual mode modulation can be achieved easily with liquid crystal spatial light modulators.[15] Figure 12 shows how a liquid crystal spatial light modulator with gray scale capability can be modified to produce dark pixels, or light pixels with a zero or π phase shift. This technique was implemented in our current experiments using a Sharp XG1500 projection television system with the only obvious drawback being the apparent reduction of LCD contrast. The use of phase coding does require the highest possible contrast from the LCD device.

3.3 Experimental Results

The experiments with our Neural Network system have demonstrated the key working capabilities expected. Examples of how the system performed in one test series are shown in the next several figures.

For the test results shown in the following figures the words "SPARTA" and "AFOSR" were used to modulate each 2D input plane. These two patterns were recorded holographically with respect to each other at each of nine wavelengths. The left side of Figure 13 shows the

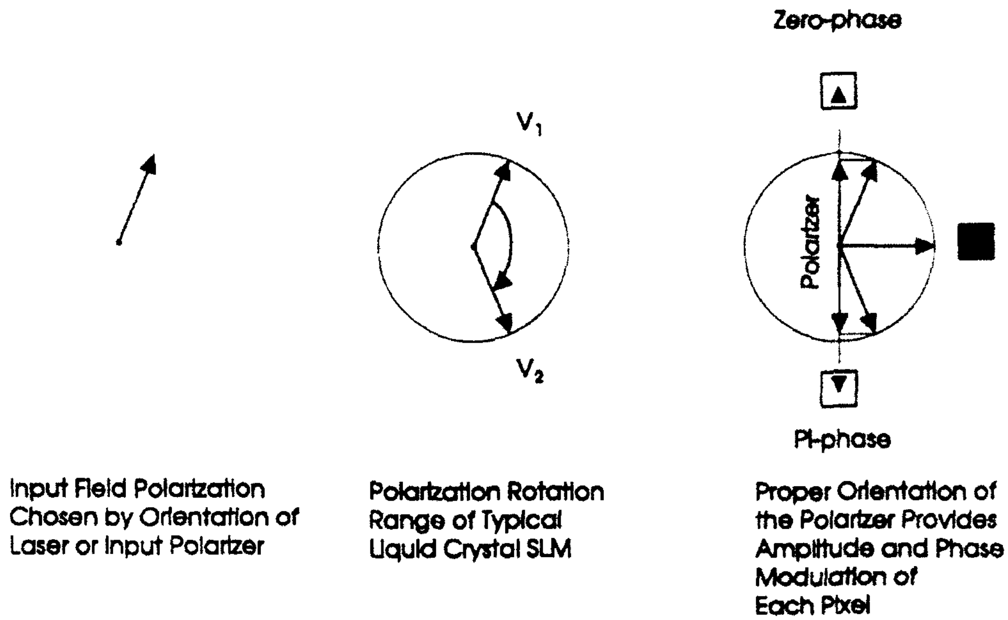


Figure 12. Reorientation of the polarizer associated with a liquid crystal SLM can produce dark pixels or bright pixels with zero or π phase shift.

binary image of the pattern presented to the SLM on one leg of the neural network. When the input pattern "SPARTA" is presented to the system and exposed at all the laser wavelengths, the image which is recalled is shown on the camera corresponding to the other leg. The image corresponding to "AFOSR" is clearly readable in the top image.

To demonstrate that the system is sensitive to changes in the input image the next two pairs of images show what happens when the system is presented with a shifted image of "SPARTA" in the middle pair of images and an image of "EOT" in the bottom pair of images. In both cases one observes only a very weak (almost zero) correspondence, as expected. (The shift sensitivity noted is a required property for the neural network and is caused by the fact that the phase code is not shifted with the input pattern.) At this point it can be clearly seen that the two input planes are properly connected and that the system is not strongly responsive to incorrect inputs.

Other tests have shown that pattern shifts as small as 1 LCD pixel in the input can provide an almost total loss of output image, showing the expected independence of the individual input nodes. Similarly, changing the input patterns to the opposite signs (i.e. from 1's to -1's and vice versa) again provides a loss of output image. These attributes are required for proper operation of the system with independent input nodes.

3.3.1 Neural Network Operation

One type of neural network is a bi directional associative memory (BAM). To test the ability of our system to operate as a BAM we begin by presenting a fragment of one of the the input images to the system. On the left side of Figure 10 the image "PART" is presented to the system, and the recalled pattern corresponding to a somewhat fragmented "AFOSR" is recalled. After being thresholded, the resulting fragmented "AFOSR" image is fed back to the system and a somewhat fragmented "SPARTA" image is recalled. Typically BAM systems are iterated for

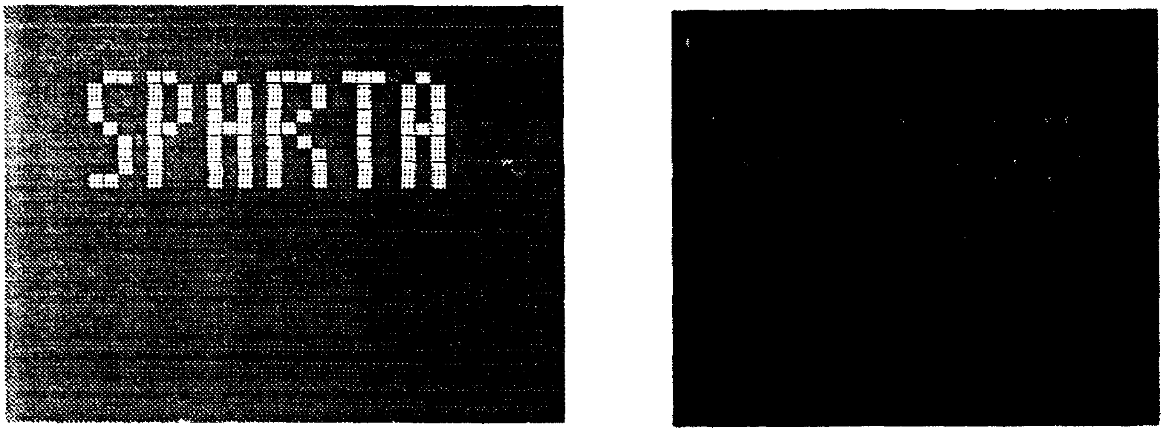
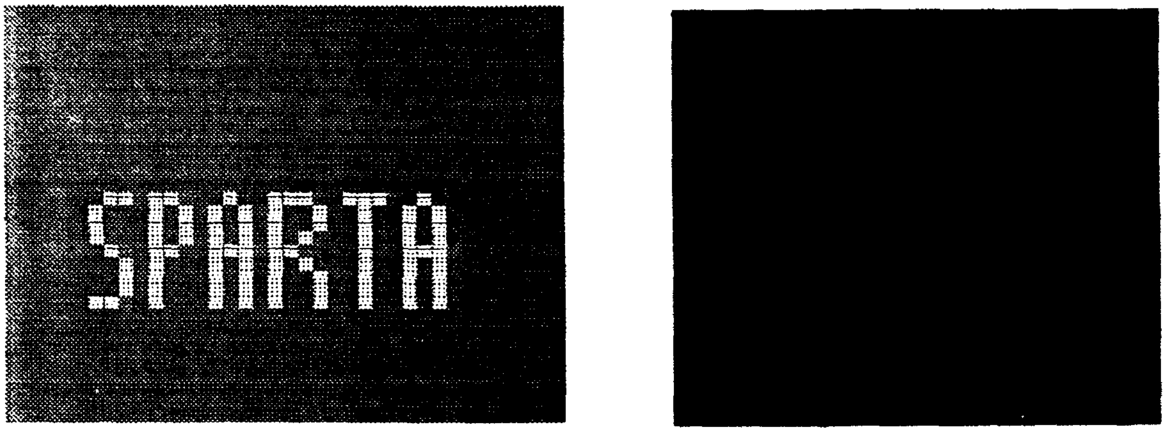
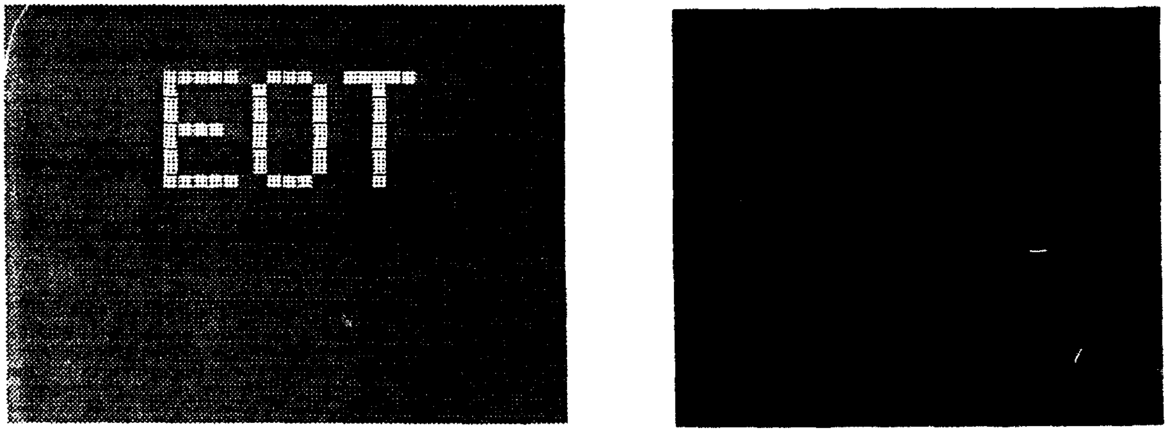


Figure 13. a. System Response to an identical input.



b. System Response to a shifted input.



c. System Response to a different input.

optimal results, therefore we chose to carry out a second iteration which presents the somewhat fragmented "SPARTA" fed forward into the system, obtaining an improved "AFOSR", which when fed back to the system again provides a somewhat improved "SPARTA". We therefore see that the 4D neural network system operates well as a BAM system. We also expect that because of its very general architecture our system can be iterated in a manner consistent with other neural networks.

At present no sophistication has been implemented such as utilizing sigmoidal filters on the CCD output since only a simple thresholding operation has been implemented so far. We expect that our system operation may be significantly improved by properly implementing this option.

3.3.2 Further Tests

Further tests have also been performed which demonstrate that the system successfully operated at close to full capacity. In the above mentioned test, only nine laser wavelengths were used, accessing only nine of the input rows of the chromatically steered input plane. In further tests it was shown that complete access to all 83 rows at all 83 wavelengths was achieved with good recording fidelity and readout. This verified the connectability of all 5,400,000 trinary input nodes.

Tests of the system as a BAM were scaled up to include more of the input nodes from each input plane. These tests scaled up operation to include 28 laser wavelengths and three separate recorded input patterns. The input patterns matched "SPARTA" with "AFOSR", "EOT" with "SHB" and "4DNET" with "NEURAL". The pattern "NULL" was used to test for a lack of correspondence to any of the recorded patterns.

The tests showed that operational performance as a BAM was observed and that all 1,800,000 trinary input nodes were operating as expected. The lack of correspondence of "NULL" to any recorded pattern was verified. Also, input the the complete "SPARTA" image resulted in the appropriate output "AFOSR" image.

The test of the system as a BAM again involved the input of "PART" and a suitable response of "AFOSR" was observed and decoded by the simple thresholding process. However, as the number of recorded input patterns has been increased, the output images became significantly noisier and more difficult to interpret. Upon feeding back the obtained "AFOSR" image into the system the proper response of "SPARTA" was observed, but not easily and cleanly thresholded by our system software. We attribute this limitation of performance as being presently traceable to several factors which we have noted earlier such as particle scattering in the sample, the limited contrast available from the active matrix LCD and the lack of sophistication in our simple thresholding software. Moreover, it is commonly known that neural networks will typically have greater difficulty making unique identifications as the number of learned associations is increased. However, successful operation as a BAM was observed with 1,800,000 interconnects operation and three pairs of unique associations recorded. Further system tests at higher capacity were not possible within the limited effort remaining in the program.

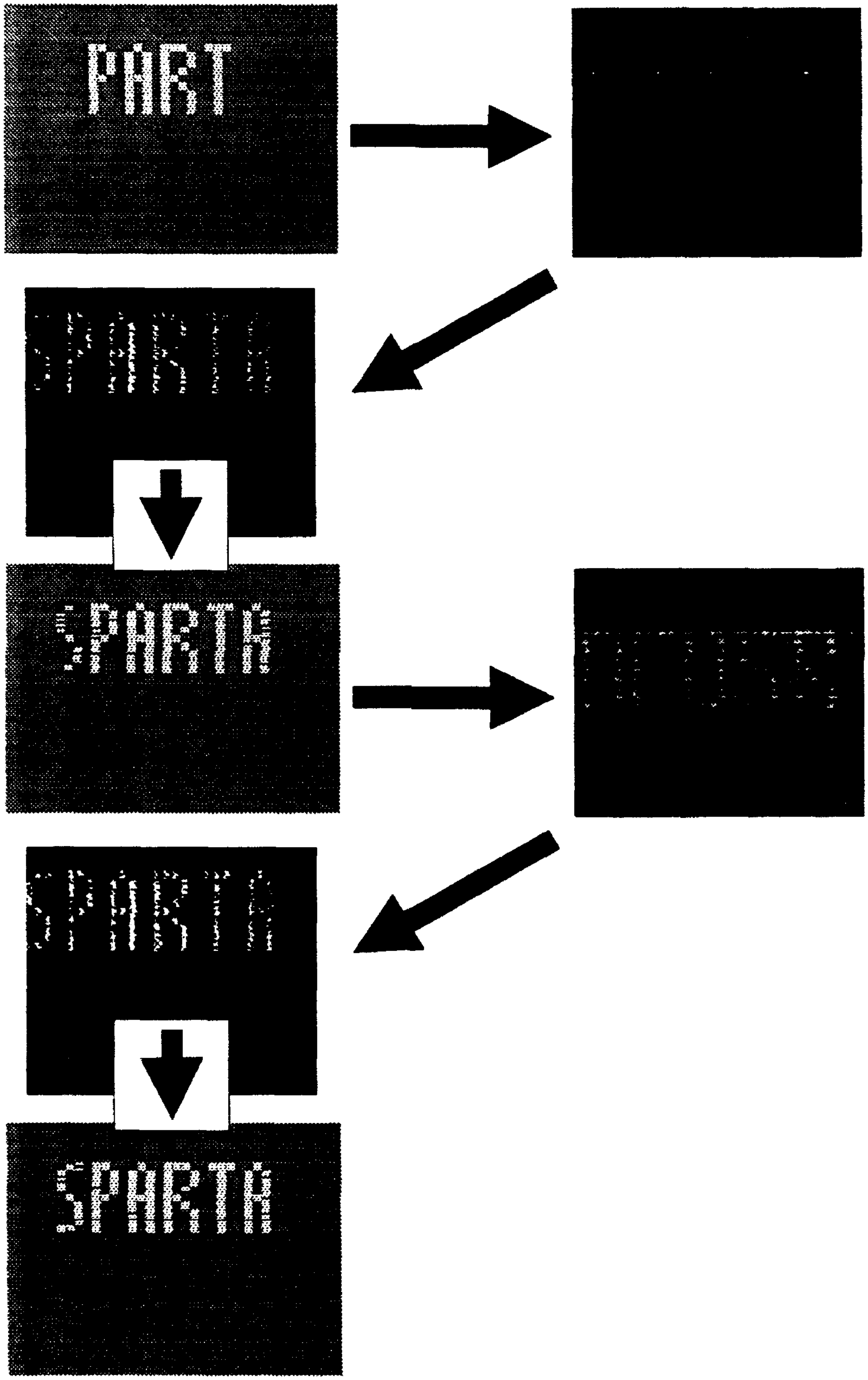


Figure 14. Neural network operation as a BAM.

4 Technology Transfer Planning

In order to examine possible commercial uses of this system, we needed to accomplish three tasks: examine potential applications of ultra-large neural networks, assess the competition, and perform a preliminary design of a next generation system. In this section, we describe the results of our analysis of potential applications, other implementations of neural networks, and our design of a next generation system with meaningful capability.

During the course of this Phase II program, we interviewed Government personnel at ONR, DARPA, AFOSR, and other DoD agencies. In addition, we attended the 2nd Government Neural Network Applications Workshop.[16] We also interviewed SPARTA experts in several areas, including BM/C3 and rotating machinery health monitoring. Finally, we examined the literature for neural network applications. Rather than attribute specific views to specific Government personnel, we have distilled these interview results by topic. Our observations about possible application areas follow.

4.1 Applications Analysis

During the course of our optical neural network demonstration project, we have demonstrated a network with over 1,800,000 interconnections (expandable to over 5,400,000 interconnections). Future versions of this architecture may be able to realize networks as large as 10^{12} interconnections. Networks of this size are larger than either electronic neural networks or digital simulations. Thus, it would appear that current work in neural network applications is limited by the availability of hardware. In fact, the current approach to neural network applications has been to find clever methods of preprocessing or to restrict the problem in order to attack real-world problems with more modest-sized networks.

In this section, we address the issue of which problems will require networks of very large size. We begin with some general observations and then proceed to specific examples.

4.1.1 Characteristics of Our 4D Optical Neural Network

Our 4D optical neural network provides a general architecture with the ability to implement networks with feedback, multilayer networks, and bidirectional networks. Learning methods such as outer product, backpropagation, and genetic algorithms can all be implemented. All of the neural network techniques which have been developed using computer simulations or electronic neural network chips can be extended to very large networks using our 4D optical neural network. Our 4D optical neural network provides the ability to learn continuously by restoring recording centers to their unrecorded state. This can be accomplished by partial photo-induced erasure using light at a much shorter wavelength (i.e. in the green). This type of "controlled forgetting" has been shown to be important for neural network operation.[17]

Finally, our 4D optical neural network provides the ability to read out the network to save all the interconnect values and to later restore the network to a previously recorded state.

4.1.2 When to Use Neural Processing

Neural networks have evolved in the animal world to make decisions when real-time performance is desired ("fight or flee"). Because in many cases a quick decision is preferable to the "best" decision, fuzzy operation is allowable and even desirable. Furthermore, rapid responses are required to inputs which have never been seen before. Our goal is to train networks using a small training set, and then apply the networks to a much larger set of inputs. One way to achieve this goal is to require similar responses to similar inputs.

In many cases, signals which exhibit marked but insignificant differences in one domain will appear quite similar in another domain. One obvious example is the power spectral density of two signals which differ only in phase. Although the time domain values of the signals will be different at almost all times, the power spectral densities will be identical. Transformation of the inputs to a neural network into a form in which the important differences are emphasized and trivial differences are de-emphasized is an important first processing step. Pre-processing using "hard-wired" systems of neurons is used in the animal world, and some types of preprocessing such as Fourier transforms can easily be implemented electronically using digital signal processors or array processors.

4.1.2.1 Pre- and Post-Processing

Some types of preprocessing are designed primarily to place the signal energy in invariant positions in the input plane. For example, spectrograms of acoustic signals are calculated by taking the magnitude of the Fourier transform over a sliding window and graphing the resulting frequency spectrum versus time. Signals with similar frequency content will thus have very similar spectrograms. Homomorphic filtering[18] can achieve the same goal for signals with other types of similarities. If a spectrogram is presented directly to a neural network, then a large number of neurons may be required to process the signal. Thus, this type of preprocessing would not reduce the requirement for a large neural network.

Other types of preprocessing may reduce the requirement for a large number of neurons. For example, studies of facial characteristics have developed a number of parameters which can be measured to determine the identity of a face or even the emotion shown by the facial expression.[19] In this case, a much smaller neural network, or even an expert system may be able to learn to distinguish faces or even emotions. Extensive preprocessing to reduce the size of the required network is a replacement for training a large network which makes use of the ingenuity of the neural network designer. Areas which have been studied sufficiently to yield a small list of significant characteristics seem to include primarily functions already performed by humans such as speech and facial recognition. This technique requires a great deal of study and knowledge, and may not be applicable to a broad range of problems.

Finally, the issue of data representation is closely related to the idea of preprocessing. J. Anderson of Brown University recommends that we study data representations used in the human brain to provide important clues to how data can be represented in an artificial neural network.[20] Anderson pointed out one data representation that is particularly relevant to the issue of neural network size. In this representation, features which can vary in intensity are represented by a sequence of neurons which conceptually resemble a liquid crystal thermometer. Low values of the feature activate one end of the sequence, while high values activate the other end of the sequence, with a sliding scale of representations in between. In this representation, many neurons are used to represent one feature value. While this observation does not mean that very large networks are absolutely required, we may find that this data representation method, evolved in the brain over millions of years, provides unexpected benefits.

Segmentation of problems in a network may allow a combination of smaller networks to be used to solve a given problem. Of course, the problem domain must be well enough understood to break the problem up prior to training. In other cases, multiple layer networks are required to solve a problem. In each of these cases, a number of smaller networks, each implementing one layer of the larger network, rather than one large network could be used to solve a problem. Thus, some problems which might initially seem to require a very large optical network can be separated into smaller problems, each of which could be solved with an electronic neural network.

4.1.2.2 Learning

One of the most important characteristics of a neural network is the ability to "learn" how to perform desired functions in a given domain without explicit programming. In fact, networks have been designed which are self-organizing, in other words, which can develop internal representations of problems without the use of training sets.[21] What are the implications for learning when using a very large network? One of the more popular methods for training multilayer networks, backpropagation, is well known for having very long training times. This technique may not be the best choice for training extremely large neural networks.

Associative memory systems may not have to deal with the problem of finding efficient training algorithms. In an associative memory system, pairs of "memories" are presented to the system. By recording the "outer product" of the two memories, the connections required to reconstruct one memory from the other are automatically established.

"Genetic algorithms" may be another way to attack this learning problem. Genetic algorithms were first developed by John Holland at the University of Michigan as a method of solving difficult problems by copying the phenomena observed in natural genetics for adapting a population to an environmental niche.[22,23] Genetic algorithm techniques have recently been used to train neural networks.[24] In these experiments, a metric was established to evaluate the network performance, and multiple networks with different weights were evaluated using this metric. Initially, the weights were chosen at random. In subsequent generations, new sets of networks were derived from modifications of the better performing networks from the previous generation. To attack complex problems, simpler or less constrained problems had to be solved first, and then additional complexity or constraints were added.

4.1.3 What's the Competition?

DARPA, in a 1988 study, estimated levels of performance which could be reached using various processing technologies. Although it is now nearly five years later, these estimates still appear relevant. Figure 15 compares projected performance of our network, in terms of the number of interconnects and interconnects per second which can be implemented, with other electronic and optical neural network implementations.[25] In our Phase I report, we related the projected performance of our neural network to the physical characteristics of our spectral hole burning recording medium and the performance of available key system elements, such as the input SLMs and output detector arrays. In addition, we showed why our 4D interconnect method

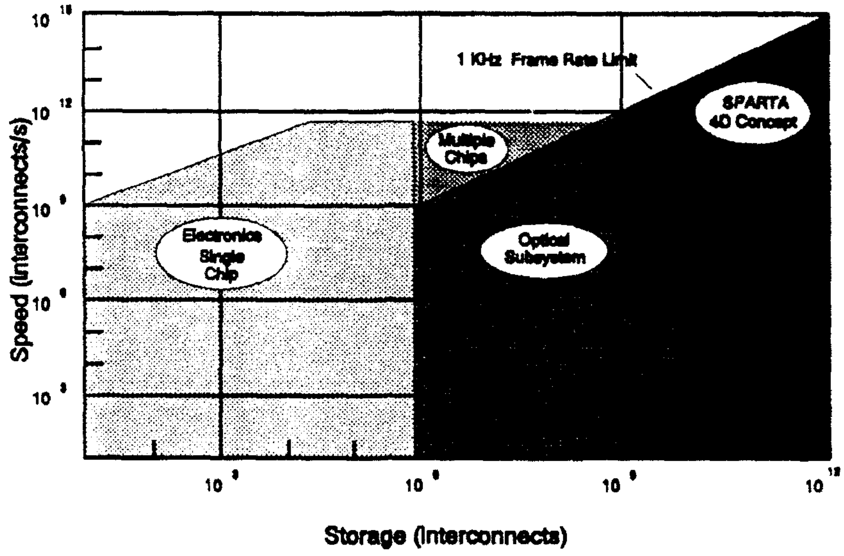


Figure 15. Projected capabilities of electronic and optical implementations of neural networks. (Adapted from DARPA Neural Network Study Final Report, 1988.)

allows full use of the space-bandwidth product of these components, unlike other proposed neural network methods.26]

Beyond other optical neural network implementations, there are three possible areas of competition for our optical neural network: human beings, electronic neural network chips, and perhaps expert systems. Human beings (and many higher animals) have extremely large networks, and can train themselves (or be trained) to perform many different types of tasks. However, humans need practical experience to be trainable. Furthermore, humans have relatively high ongoing costs, and often make mistakes due to boredom, attention span, and other factors.

The characteristic which distinguishes our optical network from electronic networks is the potentially very large size of our neural network. The largest electronic neural network we have heard of is a chip produced by E-Metrics which may be extendable to 10^8 neurons and 10^{10} interconnects per second. Although this chip cannot produce a single interconnected network of the size which can be realized using optics, in many problem areas a number of small networks can be used to solve subproblems and then these smaller networks can be combined to solve the entire larger problem. Phoneme recognition, discussed below, is an example of this type of problem.

Expert systems are useful when a list of features can be extracted and quantified (as a scalar value) which are sufficient to characterize a problem domain. The problem domain must be sufficiently limited and well understood for rules to be formulated. The parameters to be learned are then the thresholds for decisions to be made as part of "If...then..." rules. Special purpose LISP processors, which are more compact and rugged than initial versions of our optical neural network, can achieve rapid expert system processing.

4.1.4 Specific Applications

4.1.4.1 Associative and Content Addressable Memory

Neural networks have been proposed for implementation of content addressable memory (CAM) and associative memory. Clearly, this is an area in which a large capacity has utility. As in other types of computer memory, the more you have, the more you need. CAMs are directly applicable to many existing types of database manipulations including topical information retrieval, list and string processing, relational database queries, and language translation. In addition, the ability to perform extremely rapid content addressable recall from a very large database could result in entirely new and more powerful ways to achieve computer based reasoning and learning.[27] We have identified a specific architecture for CAM, which provides advantages over a more general neural network for content addressable memory.[28] The CAM we have invented, and CAMs in general, work with binary data.

Associative memory tends to work with data represented as patterns rather than data encoded as binary bits. Associative memory has a very simple training algorithm, in which pairs of memories to be associated are presented to the network simultaneously. In this way, initial experiments in associative memory could be performed which would demonstrate the characteristics and advantages of our optical neural network without introducing the additional complication of simultaneously testing training algorithms.

4.1.4.2 Resource Allocation

Resource allocation is an important problem in many areas: assignment of weapons to targets, assignment of sensors to objects, and even investment strategy. Resource allocation can be formulated in terms of the well-known traveling salesman problem (TSP).[29] The most straightforward neural network solutions for TSP only solve the 4 by 4 problem (for example, four weapons and four targets).[30,31] Genetic algorithms may be able to find good TSP solutions for larger problems. Our neural network would provide the ability to run a large population of resource allocation "solutions" in parallel to rapidly converge on a near-optimum solution.

The Hopfield formulation of the TSP requires a metric for the cost of each possible allocation - this metric may be the computationally intensive part of the problem.

4.1.4.3 Speech

Speech is an area which contains "Grand Challenges." For example, a translating telephone which would allow a Japanese speaker to converse naturally with an English speaker has been specifically called out as a "Grand Challenge" by the National Academy of Science.[32] Furthermore, the capability to implement a speaker-independent user interface could have important commercial and military implications. We believe there is a role for very large neural networks in the speech area.

Speech and language processing have many aspects: speech to text, text to speech, and machine translation, for example. Some of these problems can already be addressed by computer processing or small electronic neural networks; while others have proven to be very difficult for conventional computer approaches.

Problems which have been solved or appear well on the way to resolution include phoneme to speech[33], phoneme detection using small electronic neural networks[34], and even to some extent machine translation.[35] Text to phoneme conversion has been demonstrated using smaller neural networks than we are studying,[36] and it is not clear that the additional capacity of a very large optical neural network is required. In each case, the problem can either be reduced to a small number of rules through in-depth study of the problem, or the problem can be handled through conventional processing techniques such as relational databases.

Problems which have not been solved (to a level allowing practical user interfaces to be built) include phoneme to text conversion and text to phoneme conversion. These areas involve the requirement for a large memory, in particular to take care of the many special cases which are not governed by rules. For example, in the machine translation area, idioms (which do not make sense when translated word for word), and proverbs (which carry implicit meaning determined by a related story) have proven to be difficult to handle. Phoneme to text conversion involves problems with dropouts (all the informal contractions which appear in every spoken language), the lack of detectable separations between words, and individual variations in speech patterns. In English, which has incorporated so many words from other languages, a rule based approach becomes unwieldy when very high recognition rates are required. Even apart from the large number of rules required, it seems clear that humans do not understand speech using rules. Instead they simply learn to recognize all the many different speech patterns that they hear, including the ability to deal with different accents as they are exposed to them. Phoneme to text conversion thus seems an ideal application for a very large capacity associative memory. Text to phoneme conversion presents similar difficulties with the one exception that the source material is usually free of idiosyncrasies. (A better comparison would be handwriting to speech conversion, in which the input from the character recognition process would be full of errors and dropouts.)

Phoneme to text might be an interesting choice for a first set of associative memory experiments. There are commercial and military applications for user friendly computer interfaces, for example, natural language decision aids. Small initial experiments could be performed, followed by larger experiments. (Ultimately, a very large network will be required to store a significant number of phoneme/word pairs.) Finally, a large data base of phoneme/text information is readily available.[37–40]

In the remainder of this section, we will outline how a phoneme to text demonstration, which is extendable to a useful high performance system, could be set up.

We believe the most important requirement is for a large capacity system, which implies a large number of interconnects. This requirement implies that the input and output be spread over a large number of neurons. Spreading the input and output over a large number of neurons will determine the form of the pre- and post-processing which will be required for this application.

Although the human vocal tract can make many different sounds, English is composed of approximately 54 phonemes, which includes symbols related to punctuation. Assigning an ASCII symbol to each phoneme would clearly result in a very small number of neurons being required for a short (several second) speech sample. Similarly, representing the output text as a string of ASCII characters will not involve enough output neurons. (40,000 neurons, at two neurons per bit, can represent over 2000 eight-bit characters. We do not want to require that our network processes the equivalent of an entire page of text at once.) Furthermore, we believe that the input and output neurons to the network should be arranged so that similar inputs produce similar outputs. A binary representation does not accomplish this goal.

A better representation of input phonemes would be to create patterns, corresponding to a sequence of phonemes. One representation is shown in Figure 16. In this case, a separate neuron is assigned to each phoneme, with each column representing a time slot (of varying length to accommodate different speech patterns and speeds). Furthermore, phonemes indicating similar sounds can be grouped together to provide similar input patterns for similar strings of phonemes. Kohonen has developed two-dimensional phoneme maps using self-organizing neural networks in which similar phonemes are placed adjacent to each other.[41] Kohonen's 2D phoneme maps are similar to those created using two formant frequencies to map and identify phonemes.

Using either a phoneme versus time or a 2D phoneme representation, we believe we can create unique patterns for each string of phonemes. These patterns would have the property that similar strings of input phonemes would create similar patterns. We believe, based on our experiments with recovering an incomplete input pattern using a bidirectional associative memory approach (described earlier in this report), that this similarity of input patterns will provide error tolerance and the ability to deal with ambiguity.

The output representation should exhibit similar characteristics. A representation for output characters is shown in Figure 17. Again, each character will be represented by a separate neuron. Characters can be grouped by sound, so that, for example, "t", "p", and "b" will be placed together. Characters, such as "c", which have multiple sounds, can have two neurons devoted to them. In addition, we can have neurons devoted to combinations of characters which appear together often, such as "ch" and "sh". Studies of English have estimated that there are from 102 to 170 graphemes or basic letter groupings in English. Our 4D optical neural network could easily represent all of these graphemes, including repetitions for multiple sounds, in each output column.

In operation, a string of phonemes will slide through the input plane. Past and future phonemes will provide context for elimination of ambiguity, recognition of general speech patterns, and dealing with the dropouts associated with common speech patterns. The short delay implied by the length of the sliding window will require study to determine what is acceptable and necessary. Many classification problems require multilayer networks to solve them. However, devoting a number of neurons in an internal layer to recognizing general speech patterns or even individual speakers might be valuable.

Selection of the phoneme versus time approach discussed here or the 2D phoneme map approach mentioned above will depend on several issues. Our goal is to produce patterns for each input utterance which are relatively independent of the speaker, the speed of speech, the accent, and dropouts due to verbal contractions. A 2D phoneme map approaches this ideal in many ways. The patterns which are produced, however, do not distinguish between utterances with the same phonemes in a different order. The phoneme versus time approach, with a sliding window, clearly satisfies this requirement, but a dropout in the middle of an utterance can change a pattern in ways that would require that a network memorize multiple versions of the same general pattern, corresponding to dropouts of different phonemes. We need to examine this problem further before performing the detailed design of a demonstration.

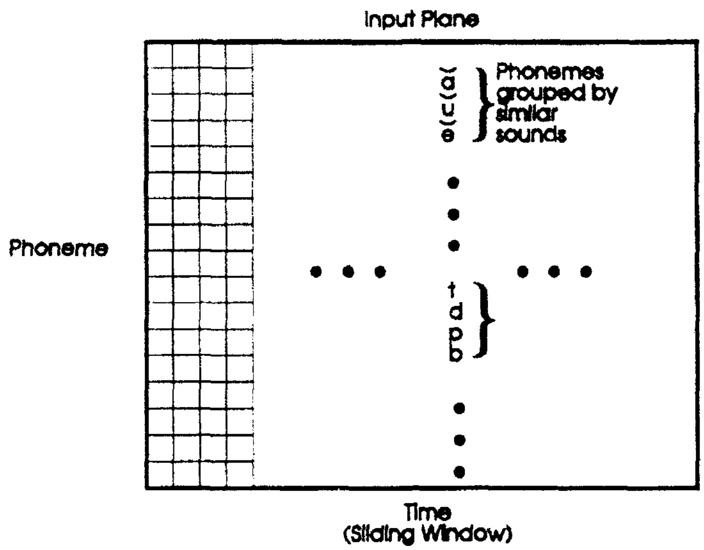


Figure 16. Representation of input phonemes.

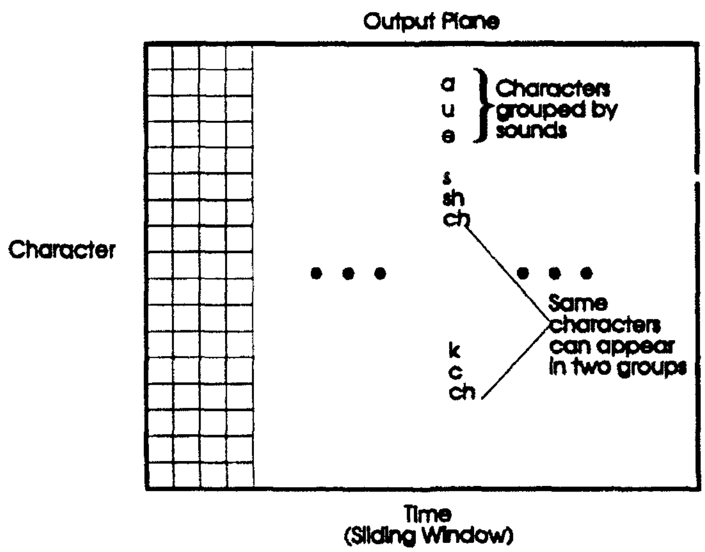


Figure 17. Representation of output characters.

4.1.4.4 Rotating Machinery Health Monitoring

Health monitoring has important military applications, for example, in the areas of helicopter maintenance and shuttle main engine turbine diagnostics.[42,43] In each case, failures during use are disastrous. This area has other attractive features, for example, Navy ships, which also have a need for health monitoring equipment, provide a large platform which would not be taxed by the space and power requirements for our 4D optical neural network.

The inputs to a neural network to perform health monitoring would be the phases and amplitudes of up to 10 harmonics of the fundamental engine rotation frequency. For example, too many high harmonics might indicate an imbalance in the machinery, while a change in the phases of the harmonics might indicate wear in gear trains, for example. Because our network has large capacity, and can combine information from widely different sources, we can also present other information to the network, such as the type of machinery, the interval since last servicing, etc.

4.1.4.5 Vision and Image Processing Applications

Vision applications may be the toughest familiar application. A large fraction of the human brain is devoted to vision processing. Furthermore, we do not understand the human visual system, beyond the most primitive levels of preprocessing. An excellent discussion of this preprocessing is given in Marr's book.[44] One ray of hope is provided by the fact that much of this preprocessing is hardwired and can also be computed efficiently with special purpose digital hardware. The key to high level vision processing will be suitable choices for the preprocessing and data representations. Because of these difficulties, we believe that vision and image processing problems are not the best choice for initial application of our neural network – they are just too difficult. Nevertheless, we will discuss a few possible applications briefly here.

Face recognition is a task which has important security applications, including building and computer system access. It is also an area is which there will be a great deal of competition from human beings. Furthermore, disguises, which can fool humans, will almost certainly fool initial neural network solutions to this problem.

This may also be an area in which feature extraction provides the ability to solve the problem with much smaller networks. Extensive study of the face has revealed a number of key features which can be used for recognition with a relatively small number of neurons and preprocessing.[45]

4.1.4.6 Radar Target Identification

Radar "images" are essentially one-dimensional projections of the outline of an illuminated target. These images tend to resemble the outline of the shape of an object. In many applications clutter can be suppressed by Doppler preprocessing, making the outline of the object less ambiguous, with fewer dropouts. Thus radar processing may be somewhat easier than the general vision problem.

Typical target recognition algorithms start by finding the outline of an object and taking the Hough transform, which graphs the line segments in the object outline by length and orientation. The Hough transform is a one-dimensional function, which can be represented on a two-dimensional plane as a sliding scale for each value of line segment orientation. For a small target set, each of these two dimensional patterns can be associated uniquely with a target. Our network could be used as a BAM to associate these two-dimensional patterns with a target type.

4.1.5 Underestimating the Size of Problems

Is it possible that all problems can be solved without recourse to large scale neural network computers such as the kind which we project are possible? This seems unlikely. While it has been shown that many problems in pattern recognition or speech recognition can be initially approached using modest sized systems, in general the error rate is unacceptably high or the problem has been scaled down to a test scenario which is artificially simplistic. It already seems clear that the higher the capacity of the system, the lower the error rate of the results and the more complex the potential outcome.

But then how large must a system be to solve any problem? This is like asking how smart one must be to solve an unsolved problem. Clearly the merit of a high capacity system can be seen, particularly when we address tasks which are aimed at mimicking and competing with human functions such as vision and speech. However, another approach can be taken which is to utilize the highest capacity system available to initially solve the problem and then "distill the solution" down to the essential components which can then be handled by an efficiently designed system of modest proportions with special purpose hardware. This is akin to the concept of the researcher who is also a teacher. The "brilliant" researcher can solve a previously unsolved problem which seemed beyond the capacity of others. Once the solution is obtained, the solution is often simple and direct enough to be understood by (i.e. taught to) many others of lesser capabilities. In this conceptual approach there will always be a need for the highest capacity system available even if the mass implementation of the solutions obtained may only require lesser hardware.

Another factor which can strongly influence the ability of a system to reliably identify correct answers is whether it is "information starved". By information starvation we mean that the information available cannot permit any system, no matter how sophisticated to identify a single unique answer. An example of this is the problem group we call optical illusions, such as interpreting 3-D features from a 2-D image. Many common 2-D images can be interpreted as being pyramids rising from a surface, or indented into a surface and no amount of analysis can distinguish the two. More information is required to solve the problem. Unfortunately, if the system capacity is very limited, providing more information may not be possible. This is where the larger capacity system can always by supported. The ability to address problems by a multiplicity of diverse inputs may be a key factor in obtaining reliable results. A humorous example of this might be for an image recognition system to identify an object as a naval minesweeper, but fail to take into account the fact that the image was obtained in a Kansas wheatfield because the system lacked the capacity to receive this input information and factor it in. From this approach we can see that the best solutions will generally be obtained by the "best and brightest" systems, meaning those having the largest capacity and which are provided with as much relevant information as possible. The principle limitation for them being implemented would then be availability, cost and suitability for the environment being addressed.

4.1.6 Summary of Applications Analysis

Large neural networks are appropriate for several types of problems:

1. problems in which we have not intensively studied the problem domain to find key features which can be used to reduce the number of neurons needed,
2. learning the key features in a problem domain by examining the structure of a large neural network which has been evolved to solve it,
3. problems requiring a large memory to store all the necessary interconnect values, and
4. problems where the data representation requires a large number of input states.

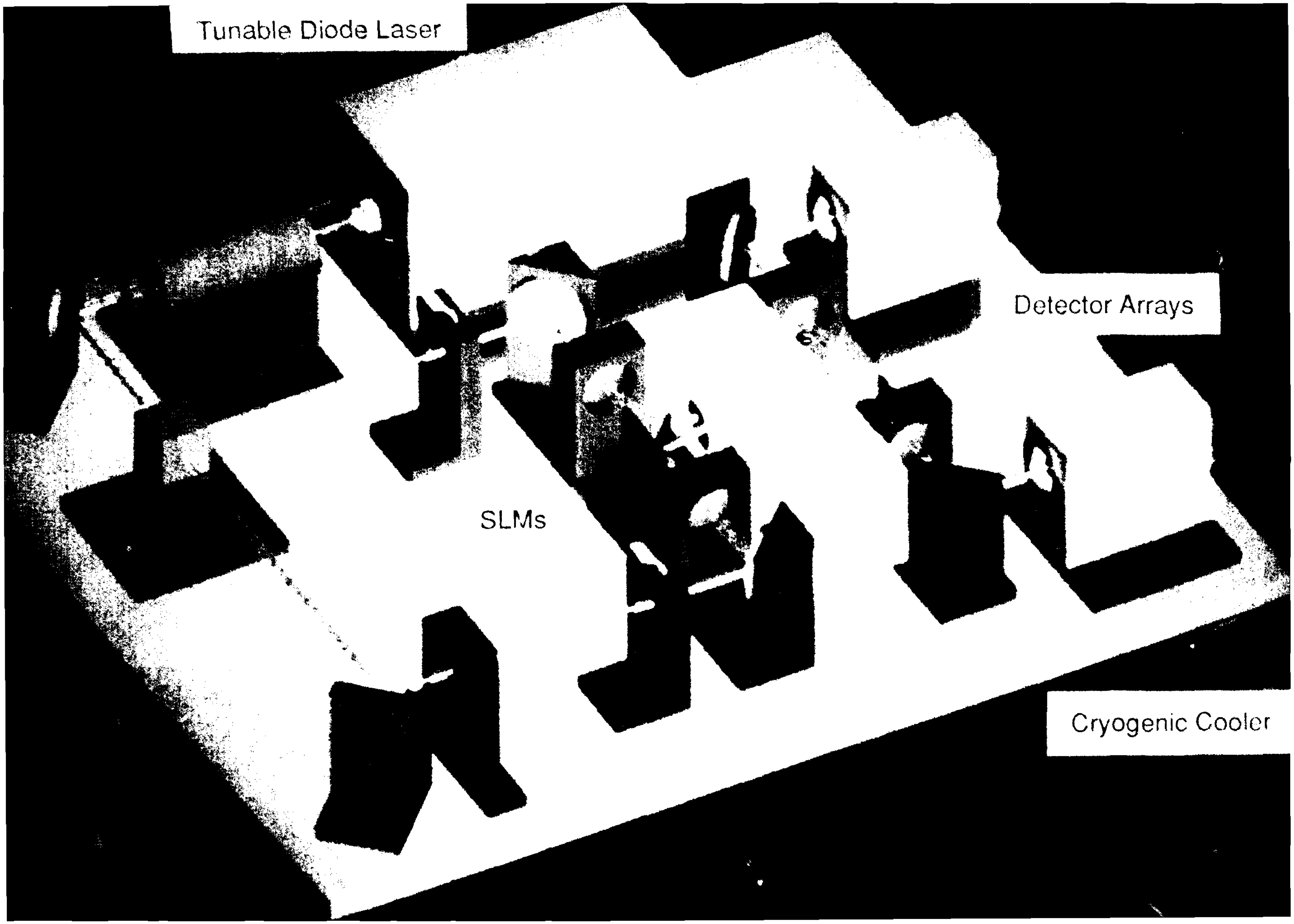
An optical neural network can be useful when preprocessing is used to place signal energy in invariant positions, but where a short list of explicit features has not been developed. This may be the case when features are not known, or when feature extraction is easily performed by the first stages of a neural network.

Specifically, accurate, high-speed, speaker-independent speech to text is an important application for the development of user friendly computer interfaces. For example, in combination with natural language query software, accurate speech to text hardware could provide an important military decision aid, and would also have important commercial applications.

4.2 SPARTA's Next Generation Neural Network System

In this section, we describe the physical configuration of a next generation neural network system. A CAD representation of this system is shown in Figure 18. The performance goals of this next generation system include 65,000 neurons, more than 4 billion interconnects, and operation at 120 billion interconnects per second. A system such as that shown in Figure 18 could be assembled using currently available or soon to be available components.

SPARTA's next generation system can be far more compact than the experimental demonstration system constructed during our Phase II effort. The optical portion of a high performance system could fit on a base approximately one half meter on a side. The recent development of visible diode lasers will permit construction of compact, tunable external cavity diode lasers in the near future, in the proper wavelength range to write in existing spectral hole burning materials. The size of the laser shown in this figure is consistent with tunable diode lasers sold by Micracor.[46] A wide variety of high speed, large format SLMs have been developed for optical computing. With the advent of HDTV, even larger format, lower cost SLMs will become available. The SLMs shown in Figure 19 are the same size as those used in our current demonstration system. Large format, low noise detector arrays are also currently available. HDTV will greatly reduce the cost of these components. High reliability, closed cycle cryostats have been developed for the semiconductor industry by CTI Cryogenics.[47] With a mean time between servicing of 10,000 hours, and a backup power supply, uninterrupted operation for a year or more can be achieved. Only the cold head is shown in Figure 18; the compressor unit, weighing approximately 100 pounds and having a volume of 2 cubic feet, is not shown in this figure. Finally, the optics required for this second generation system are all off-the-shelf items. None of these components has resolution, chromatic aberration, or field of view requirements which exceed those of typical 35 mm camera optics.


Tunable Diode Laser
Detector Arrays
SLMs
Cryogenic Cooler

5 Conclusions

The results reported here provide experimental verification of the operating principles of our neural network architecture based upon SHB media. This first demonstration system provided more than 5 million interconnects (connecting two fully populated 2D input planes each having 2324 input nodes) which is quite large by present neural network standards. All input nodes tested are fully interconnected and readily controlled entirely through a computer interface to a PC system. No factors have been encountered which are short term or long term hindrances to the continued expansion of the capabilities of this system architecture. The success of the neural network system (with all of its complexities) was possible with off-the-shelf components and points to the conclusion that sophisticated computer systems based upon SHB materials can be reliably developed.

The straightforward arguments which show how this system can be easily scaled to enormous capacities using technology which is either presently available or will soon become available in the very near future were presented in our Phase I Final Report. It is our contention that this system is scalable using the same optical principles to enormous capacities of 10^{12} interconnects or larger with operation rates reaching 10^{15} interconnects/sec.

We believe that as the data representations used in the human brain become better understood, the requirements for large neural networks and the understanding of how to use them will increase. We have identified one particular area where a very large capacity network will have utility – the real-time speaker-independent understanding of natural language. This application is clearly a dual-use technology with important military and commercial advantages.

The authors would like to express their great thanks to the consistent and timely support of Sandra Kelly in the myriad of tasks required for bringing this system into existence, and to the generous support of the Air Force Office of Scientific Research and the the Small Business Innovative Research Program. Dr. Alan Craig of AFOSR has provided guidance, encouragement, and support at every step of the way.

References

1. P. D. Henshaw, S. A. Lis, and N. R. Guivens, Jr., Final Report: Compact 4-D Optical Neural Network Architecture, Prepared for USAF, AFSC, Air Force Office of Scientific Research, Contract No. F49620-89-C-0120, Project No. 3005/A1 (April 25, 1990).
2. K. K. Rebane, R. K. Kaarli, P. M. Saari and A. K. Rebane, "Time-and-space domain holography and optical information processing based on photoburning of spectral holes," Proc. SPIE **963** (1988) 182-190.
3. U. Itoh and T. Tani, Appl. Opt. **27** (1988) 739-742.
4. H.P.H. Thijsseen and S. Volker, Chemical Physics Letters, **120** (6) 496-502 (1985).
5. S. R. Sandler and W. Karo, *Polymer Syntheses*, Academic Press, (1974) 8.
6. J. D. Roberts and M. C. Caserio, *Basic Principles of Organic Chemistry*, W. A. Benjamin, Inc., (1965) 1107-1109.
7. S. R. Sandler and W. Karo, *Polymer Syntheses*, Academic Press, (1974) 280.
8. T. Nishi, K. Arishima, H. Tabei and H. Hiratuka, Japanese Journal of Applied Physics, **27** (2) 225-228 (1988).
9. K. Sakoda, K. Kominami and M. Iwamoto, Japanese Journal of Applied Physics, **27** (7) L1304-L1306 (1988).
10. N. Kishii, S. Tamura, N. Asai, K. Kawasumi and J. Seto, SPIE Symposium: LA '89, Optical Data Storage Symposium, Post Deadline Paper-PDP1.
11. J.E. Falk, *Porphyrins and Metalloporphyrins* , Elsevier Publishing Company, (1964) 72-76.
12. W. Breinl, J. Freidrich, D. Haarer, Chem. Phys. Lett., **106**, 487 (1984).
13. D. G. Howe, H. M. Hilden and E. J. Weldon, Jr., "Error-Correcting Modulation (ECM) Codes," Proc. Topical Meeting on Optical Data Storage (March 5-7, 1990) 93-96.
14. Abdul Ahad S. Awwal, Mohammad A. Karim and Hua-Kuang Liu, "Machine parts recognition using a trinary associative memory", *Optical Engineering*, Vol. 28 (5) pp. 537-543, May 1989.
15. T. N. Tsakiris, "Electro-optic SLM-based architecture for ternary phase/amplitude filtering," Appl. Opt. **30** (1991) 4596-4598.
16. 2nd Government Neural Network Applications Workshop, 10-12 September, 1991, Huntsville, AL.
17. G. A. Carpenter and S. Grossberg, "Associative learning, adaptive pattern recognition, and cooperative-competitive decision making by neural networks," Optical and Hybrid Computing Proc. SPIE *634* (1986) 218-247.
18. A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ (1975).
19. H. Kobayashi and F. Hara, "The Recognition of Basic Facial Expressions by Neural Network," 1991 IEEE International Joint Conference on Neural Networks, Singapore, p. 460-466.
20. J. Anderson, presentation at ONR Program Review on Missile Aimpoint Selection, 10 December 1992.
21. T. Kohonen, *Self-Organization and Associative Memory*, Springer-Verlag, New York (1984).

22. J. H. Holland, *Adaptation in Natural and Artificial Systems*, U. Michigan Press, Ann Arbor (1975).

23. L. Davis, ed. *Handbook of Genetic Algorithms*, van Nostrand Reinhold, New York (1991).

24. H. de Garis, "Genetic Programming, Building Nanobrains with Genetically Programmed Neural Network Modules," 1990 IEEE International Conference on Neural Networks, San Diego, p. III-511-III-516.

25. DARPA Neural Network Study, Final Report, MIT Lincoln Laboratory (18 July, 1988).

26. P. D. Henshaw, S. A. Lis, and N. R. Guivens, Jr., "Compact 4-D Optical Neural Network Architecture," SPARTA LTR90-006 (25 April, 1990) Contract No. F49260-89-C-0120.

27. C. Stanfill and D. Waltz, "Toward Memory-Based Reasoning," Comm. of the ACM **29** (1986) 1213-1228.

28. P. D. Henshaw and S. A. Lis, "Content Addressable Optical Data Storage System," filed 17 July, 1992.

29. J. J. Hopfield and D. W. Tank, " 'Neural' Computation of Decisions in Optimization Problems," Biological Cybernetics **58** (1988) 63-70.

30. G. V. Wilson and G. S. Pawley, "On the Stability of the Traveling Salesman Problem Algorithm of Hopfield and Tank," Biological Cybernetics **58** (1988) 63-70.

31. *Neural Computing*, NeuralWare, Inc., Pittsburgh (1991) 244-255.

32. *The National Challenge in Computer Science and Technology*, National Academy Press, Washington, DC (1988) p. 34.

33. For example, the Texas Instruments Speak and Spell.

34. A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. Lang, "Phoneme Recognition Using Time-Delay Neural Networks," IEEE Trans. Acoustics Speech and Signal Processing, **ASSP-37** (1989).

35. "French Assistant," Microtac Software, 4655 Cass Street, Suite 214, San Diego, CA 92109.

36. T. J. Sejnowski and C. R. Rosenberg, "NETtalk: A Parallel Network that Learns to Read Aloud," Johns Hopkins University Technical report JHU/EECS-86/01.

37. E. Pulgram, "Phoneme and Grapheme: A Parallel," Word – J. of the Linguistic Circle of New York, 7 (1951) 15-20.

38. R. A. Hall, Jr. *Sound and Spelling in English*," (Chilton, New York) 1961.

39. E. J. Gibson, H. Osser, and A. D. Pick, "A Study of the Development of Grapheme-Phoneme Correspondences," J. of Verbal Learning and Verbal Behavior **2** (1963) 142-146.

40. L. Garvin and E. C. Trager, "The Conversion of Phonetic into Orthographic English: A Machine-Translation Approach to the Problem," Phonetica **11** (1964) 1-18.

41. T. Kohonen, *Self-Organization and Associative Memory*, (Springer-Verlag, New York) 1984, p. 148.

42. M. Tillman, J. Pooley, W. Teoh, M. Whitley, G. Lyles, and F. D. Bachtel, "Employment of Neural Network Methodology to Space Shuttle Main Engine (SSME) Health Monitoring (HM)," SPARTA, Inc. (1989).

43. M. J. Boek, "Experiments in the Application of Neural Networks to Rotating Machine Fault Diagnosis," IJCNN (1991) 769-774.

44. D. Marr, *Vision*, (Freeman, New York) 1982.

45. H. Kobayashi and F. Hara, "The recognition of Basic Facial Expressions by Neural Network," 1991 IEEE International Joint Conference on Neural Networks, pp. 460-466.

46. Micracor, Concord, MA.

47. CTI Cryogenics.